

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2005, 13:04:08 ; Search time 42 Seconds
(without alignments)
1683.793 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869

Sequence: 1 MNQNNDNVEIIDSHTSPY.....KLETIQIKINTFFNHTKTL 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: Pir1:*
- 2: Pir2:*
- 3: Pir3:*
- 4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1061.5	27.4	1180	2 A26858	parasporal crystal
2	1060.5	27.4	1180	2 I39870	parasporal crystal
3	893.5	23.1	934	2 B29838	parasporal crystal
4	758.5	19.6	1136	1 USB881	parasporal crystal
5	676	17.5	659	2 SI0228	parasporal crystal
6	654	16.9	652	2 I39811	parasporal crystal
7	641.5	16.6	1138	2 A48944	parasporal crystal
8	618	16.0	1160	2 I40589	parasporal crystal
9	612	15.8	1157	1 S49247	parasporal crystal
10	592.5	15.3	652	2 A27323	parasporal crystal
11	576	14.9	1228	2 S00873	parasporal crystal
12	568.5	14.7	719	2 I39815	parasporal crystal
13	565.5	14.6	719	2 I39814	insecticidal prote
14	565.5	14.6	719	2 S25383	insecticidal prote
15	564	14.6	649	1 JH0261	parasporal crystal
16	553.5	14.3	719	2 I40590	parasporal crystal
17	547.5	14.2	1174	2 S32649	parasporal crystal
18	544.5	14.1	1165	2 S11446	parasporal crystal
19	539.5	13.9	1178	1 USBSXH	parasporal crystal
20	528	13.6	1154	2 S39536	parasporal crystal
21	526	13.6	1177	2 A49785	parasporal crystal
22	516	13.3	1174	2 A42459	parasporal crystal
23	512.5	13.2	618	2 S11445	parasporal crystal
24	512.5	13.2	1171	2 A37829	parasporal crystal
25	512.5	13.2	1171	2 I40572	parasporal crystal
26	508.5	13.1	1176	2 A48970	parasporal crystal
27	505	13.1	1160	2 S32647	parasporal crystal
28	492.5	12.7	1166	2 S32645	parasporal crystal
29	487	12.6	655	2 JC7140	protoxin - Bacillu

30	483	12.5	1172	2 S32689	parasporal crystal
31	481	12.4	1181	2 A41052	parasporal crystal
32	480	12.4	823	2 S04181	parasporal crystal
33	476.5	12.3	934	2 A22798	parasporal crystal
34	475.5	12.3	1176	2 JT0241	parasporal crystal
35	475	12.3	1189	2 S00944	parasporal crystal
36	472	12.2	613	2 JC6033	mosquitocidal prot
37	471	12.2	1156	2 S19306	parasporal crystal
38	469.5	12.1	1176	2 JC2219	parasporal crystal
39	469	12.1	1155	2 I39838	parasporal crystal
40	465	12.0	1155	2 A26513	parasporal crystal
41	465	12.0	1155	2 JD0002	parasporal crystal
42	465	12.0	1156	2 A29125	parasporal crystal
43	464.5	12.0	1176	2 A32617	parasporal crystal
44	464.5	12.0	1176	2 S02215	parasporal crystal
45	464	12.0	1155	2 S02134	parasporal crystal

ALIGNMENTS

RESULT 1

A26858

parasporal crystal protein cry4Aa1 - Bacillus thuringiensis subsp. israelensis
N:Alternate names: Parasporal crystal protein cryIVA

C:Species: Bacillus thuringiensis subsp. israelensis

C:Date: 19-Nov-1988 #Sequence_revision 19-Nov-1988 #text_change 05-Oct-2004

C:Accession: A26858; S48691

C:Word: E.S.; Ellar, D.J.

Nucleic Acids Res 15, 7195, 1987

A:Title: Nucleotide sequence of a Bacillus thuringiensis var. israelensis gene encoding ?

A:Reference number: A26858; MUID:88015571; PMID:2821500

A:Accession: A26858

A:Molecule type: DNA

A:Residues: 1-1180 <WAR>

A:Cross-references: UNIPROT:P16480; UNIPARC:UPI000016EA42; GB:Y00423; NID:G40351; PIDN:CJ
A:Note: the authors translated the codon GCA for residue 308 as Thr

R:Nishimoto, T.; Yoshisue, H.; Ihara, K.; Sakai, H.; Komano, T.

FEBS Lett. 348, 249-254, 1994

A:Title: Functional analysis of block 5, one of the highly conserved amino acid sequences

A:Reference number: S48691; MUID:94307434; PMID:7913448

A:Accession: S48691

A:Molecule type: DNA

A:Residues: 667-676 <NIS>

A:Cross-references: UNIPARC:UPI000017819A

C:Superfamily: Parasporal crystal protein

C:Keywords: delta-endotoxin

Query Match 27.4%; Score 1061.5; DB 2; Length 1180;

Best Local Similarity 35.1%; Pred. No. 7e-62;

Matches 271; Conservative 118; Mismatches 278; Indels 105; Gaps 24;

Qy	1	MNQNDNNEVEIIDSHTSPYFPFRNSDSRVPYTNPNQPLQNTNYKEWLNCOGNTQYG	60
Db	1	MNPYQNKVEYETLNASQKKL--NISNNTRYPIENSPKQLQSTNYKDWLNWCQQQYQYG	58
Qy	61	DNFTFASADTIAAVSAGTIIVSGTLGGLTISGPIIGTIIISFGTLITVFPAG	120
Db	59	GDFHTFDS---GELSAITIVGVTLTGFGFTT----PLGL---ALIGGTILIPVLPFAQ	108
Qy	121	EQDKTVMQFIMGEIFVDTPLTESIKQLKLTLEGFRHQILQSYNTALDDWRKLKLQAP	180
Db	109	DQSNIT--WSDFTQTKNIIEKIASTIYSNANKILNRSFNVISTYHNHLKTWE-----NNP	162
Qy	181	GLPSSALQQAALTUKIRFENVHNDIFREIP--GFQLEYTKTLLPIYAAANFHLNLO	238
Db	163	NPQNTQDVRTQIQLVHVFQNVIPVLVNSCPNPSPDCDYINILVLSYAAANLHLTVLN	222
Qy	239	QGAELADEWNADIHPSQIEPNAGTSDDYVKLLKENIPKYSNYCANTYRGLKNLRDEP--	296
Db	223	QAVKFEAYLKNRQFDYLEP--LPTAIDYYPVLTKAIEDYTNVCVITYKKGLNLKTTPOS	281
Qy	297	----NMKWSIFNDYRRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKIGIKNELTREIYTT	352

Db 282 NLDGNINWNTYRTKMTTAVLDVVALFVNDYVGKY-----PIGVQSELTRREIYQV 333
QY 353 EINFDRLPQLRVQPNLATWEYNLTRASFPLFSLEQFIY-----TENTNFGNR----- 401
Db 334 -LNFESPYKY--DFQYQEDSLTRRP-HLFTWLDLSLNFYEKAQTTPNFFSHYNNMFHY 389
QY 402 -LVGISNRDAPYTSNITETLYGERTGSPPTKTRPFESYKVSIVTDROQSPVSPQPHF 460
Db 390 TLNDSIQKSVFGNHNVTDL--KSLGLATNIYI-----F 422
QY 461 IINOIEL---YLNGSSNNT-----LKYSAGSGLSNYO-NTTFFQPPRKDC 502
Db 423 LLNVISLDNKLNDYNNISKMDFFITNGRLLEKELTAGSGQITVDNKNIFGLPILKRR 482
QY 503 NLVIDGCSFNFNYSHILSHFSLFYYSYVIGLQLDITGLVGTWTHSSVDYRNAISDKI 562
Db 483 ENQGNPTLFTPDYNYSHILSFKLSIPATYKTQVY-----TFAWTHSSVDPKNTIYTHL 537
QY 563 ITWIPAKGNLDTNSKVIIEGPGHGTGNLVYLOSQRLEITCTETPNSTQSYFIRLRYATN 622
Db 538 TTQIPAVKANSLGTASKVQVQPGHGTGDLI--DFKDHFKITCQHSNFFQSYFIRIRYASN 595
QY 623 GAGNTLPNLSLITPGVIGIPQRLNNTFSGTNNYNNLQYGDYGFYQFPSTVTLPLNRPNIPF 682
Db 596 GSANTRAVINLSIPGVAEL--GMALNPTFSGTDYTNLKYDFQYLFESNEVFPAPNQNISL 654
QY 683 IFNRADV--SNSILLIDKIEFIPITSSMHQNRKQKLETIQTKINTFTNHTK 733
Db 655 VFNRSDVYNTTTLIDKIEFLPITRSIREDRKQKLETVQOIINTFYANPIK 706
RESULT 2
I39870
parasporal crystal protein ISRH4 - Bacillus thuringiensis
C:Species: Bacillus thuringiensis
C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 05-Oct-2004
C:Accession: I39870
R:Sen, K.; Honda, G.; Koyama, N.; Nishida, M.; Neki, A.; Sakai, H.; Himeno, M.; Kom. no, Agric. Biol. Chem. 52, 873-878, 1988
A:Title: Cloning and nucleotide sequences of the two 130 kDa insecticidal protein genes
A:Reference number: I39869
A:Accession: I39870
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1180 <RES>
A:Cross-references: UNIPROT:P16480; UNIPARC:UPI0000001B1F; GB:D00248; NID:g216289; I IDN:
C:Superfamily: Parasporal crystal protein
C:Keywords: delta-endotoxin
Query Match 27.4%; Score 1060.5; DB 2; Length 1180;
Best Local Similarity 35.1%; Pred. No. 81e-62;
Matches 271; Conservative 118; Mismatches 278; Indels 105; Gaps 24;
QY 1 MNQNDNNYEIIDSHTSPYFPNRSNDRSRYPTNNPNQPLQNTNYKWLNMCCQNTQYG 60
Db 1 MNPYQNKNEYELNASQKLL--NISNNYTRYPIENSPKQLQSTNYKDWLNMCCQNTQYG 58
QY 61 DNFETPASADTTAAVSAGTIVSGTLLAGIGLTSISGPIIGIIGAILISFGTLITVFWPAG 120
Db 59 GDFETPIDS--GELSAITYIVGTVLTGFGFTT-----PLGL--ALTGFTGLIPVLPFAQ 108
QY 121 EODKTVMTQFIKWEI FVDTPLTESIKQLKLTLEGFRQILQSNTALDWRKLRLOAP 180
Db 109 DOSNT--WSDFITQTKNIKKBIASITYSNANKILNRSFNVISTYHNLKTWE-----NNP 162
QY 181 GLPPSSALQQAALTALKIRFENVHDFIREIP--GFOLETYKTLILLPIYAQAANFHLNLLQ 238
Db 163 NPQNTQDVATQQLVHYHFNQVNPILVNSCPNPNPSCDYNNILVLSYAQAANHLTLN 222
QY 239 QGAELADEWNADHPQIEPNAGTSDDYKLLKKNIPKYSNYCANTYRTGLKNLRDEP-- 296
Db 223 QAVKFEAYLKNRQFDYLEP-LPTAIDYVPVLTAKAIEDYNTYCVTVYKKGLNLIKTPDPS 281

QY 297 ----NMKWSIFNDYRRYMTITVLDTISQPSLYDIKRYRDSIGGIEVKGIKNELTREIYTT 352
Db 282 NLDGNINWNTYRTKMTTAVLDVVALFVNDYVGKY-----PIGVQSELTRREIYQV 333
QY 353 EINFDRLPQLRVQPNLATWEYNLTRASFPLFSLEQFIY-----TENTNFGNR----- 401
Db 334 -LNFESPYKY--DFQYQEDSLTRRP-HLFTWLDLSLNFYEKAQTTPNFFSHYNNMFHY 389
QY 402 -LVGISNRDAPYTSNITETLYGERTGSPPTKTRPFESYKVSIVTDROQSPVSPQPHF 460
Db 390 TLNDSIQKSVFGNHNVTDL--KSLGLATNIYI-----F 422
QY 461 IINOIEL---YLNGSSNNT-----LKYSAGSGLSNYO-NTTFFQPPRKDC 502
Db 423 LLNVISLDNKLNDYNNISKMDFFITNGRLLEKELTAGSGQITVDNKNIFGLPILKRR 482
QY 503 NLVIDGCSFNFNYSHILSHFSLFYYSYVIGLQLDITGLVGTWTHSSVDYRNAISDKI 562
Db 483 ENQGNPTLFTPDYNYSHILSFKLSIPATYKTQVY-----TFAWTHSSVDPKNTIYTHL 537
QY 563 ITWIPAKGNLDTNSKVIIEGPGHGTGNLVYLOSQRLEITCTETPNSTQSYFIRLRYATN 622
Db 538 TTQIPAVKANSLGTASKVQVQPGHGTGDLI--DFKDHFKITCQHSNFFQSYFIRIRYASN 595
QY 623 GAGNTLPNLSLITPGVIGIPQRLNNTFSGTNNYNNLQYGDYGFYQFPSTVTLPLNRPNIPF 682
Db 596 GSANTRAVINLSIPGVAEL--GMALNPTFSGTDYTNLKYDFQYLFESNEVFPAPNQNISL 654
QY 683 IFNRADV--SNSILLIDKIEFIPITSSMHQNRKQKLETIQTKINTFTNHTK 733
Db 655 VFNRSDVYNTTTLIDKIEFLPITRSIREDRKQKLETVQOIINTFYANPIK 706
RESULT 3
B29838
parasporal crystal protein - Bacillus thuringiensis subsp. israelensis (fragment)
C:Species: Bacillus thuringiensis subsp. israelensis
C:Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 05-Oct-2004
C:Accession: B29838
R:Thorne, L.; Garduno, F.; Thompson, T.; Decker, D.; Zounes, M.; Wild, M.; Walfield, A.M. J. Bacteriol. 166, 801-811, 1986
A:Title: Structural similarity between the Lepidoptera- and Diptera-specific insecticidal A;Reference number: A94672; MUID:86223796; PMID:3011746
A:Accession: B29838
A:Molecule type: DNA
A:Residues: 1-934 <THO>
A:Cross-references: UNIPROT:P09662; UNIPARC:UPI000001781A0; GB:M12662; NID:g143228
Query Match 23.1%; Score 893.5; DB 2; Length 934;
Best Local Similarity 31.9%; Pred. No. 6.3e-51;
Matches 245; Conservative 127; Mismatches 267; Indels 129; Gaps 30;
QY 1 MNQNDNNYEIIDSHTSPYFPNRSNDRSRYPTNNPNQPLQNTNYKWLNMCCQNTQYG 60
Db 1 MNPYQNKNEYELFNAPSNGF--SKSNYSRYPLANKPNQPLKNTNYKDWLNMCCQNTQYG 58
QY 61 DNFETPASADTTAAVSAGTIVSGTLLAGIGLTSISGPIIGIIGAILISFGTLITVFWPAG 120
Db 59 NNAGNFASSETIVGVSGAGIIVGTMLG-----AFAAP--VLAAGIISFGTLPIFW--QG 109
QY 121 EODKTVMTQFIKWEI FVDTPLTES--SIKQLKLTLEGFRQILQSNTALDWRKLRKL 177
Db 110 SDPANVQDNLNG-----GRPIQEKNIIVNLTSIVTPIKNQLDKQEFDFKKEPAR-- 163
QY 178 QAPGLPSSALQQAALTALKIRFENVH--DFIREIFGFOLETYKTLILLPIYAQAANPHL 234
Db 164 -----THANAKAVHDLFTTLEPIIDKLDLKNNASYRIPT-----LPAYATATWHL 211
QY 235 NLLQGAELADEWNADHPQIEPNAGTSDDYK--LLKENIPKYSNYCANTYRTGLKNLR 293
Db 212 NLLKHAATYNNIW---LQNOGINPSTFNSNYGYLKRKIQEYTDYCIQTYNAGLTWIR 268

[illegible]

A;Reference number: I39869
A;Accession: J39869
A;Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-202, 'C', 205-271, 'L', 273-324, 'Y', 326-1136 <RES>
A;Cross-references: UNIPARC:UP1000015E4D; GBI:D00247; NID:g2i6z87; PIDN:BAA00178.l.; PIDs:
R;Angusthanasombat, C.; Chungtichaiornchai, W.; Kertbundit, S.; Luxananil, P.; Settasattar
Mol. Gen. Genet. 208, 384-389, 1987
A;Title: Cloning and expression of 130-kd mosquito-larvicidal delta-endotoxin gene of Bac
A;Reference number: I40584; MUID:8803B331; PMID:2890080
A;Accession: I40584
A;Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A;Residues: 1-142 <XES>
A;Cross-references: UNIPARC:UP1000016EA41; EMBL:X05692; NID:g40350; PIDN:CAG29174.1.; PID:
C;Comment: This protein is toxic to many lepidopteran larvae.
C;Superfamily: Parasporal crystal protein
C;Keywords: delta-endotoxin

Query Match 19.6%; Score 758.5; DB 1; Length 1136;
Best Local Similarity 31.1%; Pred. No. 7.2e-42;
Matches 232; Conservative 123; Mismatches 266; Indels 125; Gaps 33;

Qy	28 DSRYPYNPNPQNONTNYKSLWLNCCOQTGYDNFETTFASADTTIAAVSAGTIVSGTLTA 87
Dy	: : : : : : : : : :
Dy	2 NSGYPELANDQGSMKNNTKYDWLAMCNNQQGVN---PAAINSSSVSTALKVACAI-- 55
Qy	88 GIGGLTSISGPIGHIIGAILISFGTLITVFWPAG-EODKTVMTOFIKMGEIFVDTPLTESI 146
Dy	: : : : : : : : : : : : : : : : :
Dy	56 -----LKFNVPAPAGVLTVL-----SAVLPLILPWPTNPTEPERVMNDFTMTGNLDIQDTVTAY 108
Qy	147 KQLKLQPLEGRFOILOSYNTALDDWRKKLRLOAQGLPPSSALQQAALTIKIRFENVHNF 206
Dy	: : : : : : : : : : : : : : : :
Dy	109 RTANAKWTVVKDLYDOYTTKENTWK-----EPNNQSVRTAVIT---QF-NITSAR 156
Qy	207 IREIPGF-QLEYTKLLLPPIYAQAAPHMLNQGAELADEWNADIHPSQIEPNAGTSD 264
Dy	: : : : : : : : : : : : : : : : :
Dy	157 LRETAVYSFLVGVELLLLPIYAQVANFNLLIRDGLINAQEWS-----LARSAG-D 207
Qy	265 DYVKLKAENPKYSNCANYRTGLKNLRDEPNMKWSIFENDRYRRYMTITVLDTISOFSLY 324
Dy	: : : : : : : : : : : : : : : : :
Dy	208 QLVNTWMVOYTKEYIAHSITWNGKGDLVRNLSNGOWITFDNKREMTIOVLDFIALFASY 267
Qy	325 DIKRY-RDSIGGIEVKIGNELTREITYTEINFORLPOLRVQPNLATMEYNLTRASFCLF 383
Dy	: : : : : : : : : : : : : : : :
Dy	268 DPRYPADKIINTKLS-KTEFTREIYTA-----LVESPCKSIAPAEEALTR-DVHLF 318
Qy	384 SFLEQPFIFTENTNFCNRNVLCINSRDAPTYSNTIT---ETLYGER-TGSPTTKTIIRFPBS 439
Dy	: : : : : : : : : : : : : : : : :
Dy	319 TWLKRVDFT-NTIIQDLRFSLANKIGFSYTNSSAMOBEGSIGSSGFGSNLTQHQLNSN 377
Qy	440 -YKVSVTDQRSPPSPHQPHFINOLEYLINGSNNTLK---YSAGCSLSNYQN----- 490
Dy	: : : : : : : : : : : : : : : : :
Dy	378 VYKTSI-TDTSSP-----SNRVTKMDFYKIDGTLASYSNITPT 415
Qy	491 ----TFFOFFRKOCNVLDPDCSPNFNNYSHLGHSLFTVSYYVIQLOLIIDTG-- 543
Dy	: : : : : : : : : : : : : : : : :
Dy	416 PEGLRTIFFGFSNEN----TPN-QPVNDYTHILSY-----IKTDVIDYSN 458
Qy	544 --VLGWTHSSVDRYNAISKDIITWIIPAIGNLDNKSKEPGHGHTGGNLYLOS----Q 597
Dy	: : : : : : : : : : : : : : : : :
Dy	459 RVSFAMTHKIVIDPNNQIYDAITCVPAVSKNFLATAVKIKGPGHTGGDLVALTSNGTLS 518
Qy	598 GRLEICTCEP---NSTOSYFILRAYATNGANTLPNISLTIPGVIGIPQRNLNWTFSGTN 654
Dy	: : : : : : : : : : : : : : : : :
Dy	519 GRMEIOCKTSIFNDPTRSYGLRIYAAANSPP--IVLANSVYLQGVSRGTTISTESTFSREN 576
Qy	655 Y---NNLOYGDGFYQ-FPSVTVLPLARN--IPIFNRADY-SNSILLIIDKEFTPITSS 707
Dy	: : : : : : : : : : : : : : : : :
Dy	577 NIITPDULKYEERYKPDPFAIVPMRLSSNOUITAITPOLNMNTSNNOVIIIDRIEIPIPOT 736
Qy	708 MHONREQOKLETIQTKINTPEFTNHFK 733
Dy	: : : : : : : : : : : : : : : : :

Db 637 VLDETONQNLESREVVNALFTNDK 662

RESULT 5
S10228
parasporal crystal protein cry3Ba1 - Bacillus thuringiensis (fragment)
N:Alternate names: coleopteran-active parasporal crystal protein; delta-endotoxin
C:Species: Bacillus thuringiensis
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 05-Oct-2004
C:Accession: S10228
R:Sick, A.; Gaertner, F.; Wong, A.
Nucleic Acids Res. 18, 1305, 1990
A:Title: Nucleotide sequence of a coleopteran-active toxin gene from a new isolate, f Ba
A:Reference number: S10228; MUID:90206811; PMID:2320431
A:Accession: S10228
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-659 <SIC>
A:Cross-references: UNIPROT:P17969; UNIPARC:UPI0000126C13; EMBL:X17123; NID:940258; PIDN
C:Genetics:
A:Gene: cryIIIB
C:Superfamily: Parasporal crystal protein
C:Keywords: delta-endotoxin; toxin

Query Match 17.5%; Score 676; DB 2; Length 659;
Best Local Similarity 28.0%; Pred. No. 8.9e-37;
Matches 208; Conservative 116; Mismatches 286; Indels 134; Gaps 28;

QY 1 MNQNDNNEYEIIDSTSPYFNNRNSDRYPYTNPNQPLQNTYKWLWMCQNTQYG 60
DB 9 MNPNN-RSEYDTIKVTFNSELP---TNHQYPLADPNSTLELNKFEFLRMADNST-- 62

QY 61 DNFETPASADTAAVAGATVSGTLAIGLGLNSISGPIGIIAIIISFGTLITVFPAG 120
DB 63 ----EVLDSSTVKDAGTGVGVQILGVG-----VFPAGALTFFYOSFLNALWPS- 110

QY 121 EODKTVWTOFIKMGEIFVDTPLTESIKQLKLTLEGPRQLQSYNTALDWRKLRQAP 180
DB 111 --DADPWKAFMAQVEVLIDKKIEYAKSKALAELOGLQNNFEDYVNALDSWKK----- 161

QY 181 GLPPSALQQAALTAKIRENVHNDFIREIPGQLETYKTLPIYAQAANFHLNLQOG 240
DB 162 -APVNLRSRRSQDRIRLEFSQAESHFNPSFAVSKFEVLFLPTVAQAANTHLLIKDA 220

QY 241 AELADSEWADIHPSQIEPNAGTSD--YKLLKENIPKYSNYCANTYRTGLKNLRDEPN 297
DB 221 QVFGGEWY-----SSEDIAEFYQRLKLTQQYTDHCNVNWNVGLNSLRGSTY 268

QY 298 MKWSIFNDYRRVMTITVLDTISQFSLYDIKRYRDSIGGIEVKGIKNELTREIYTTTEINF 357
DB 269 DAWVKENRRERMTLVLDLIVLPFYDRLYS-----KGVKTELTRDIETDPI-FT 319

QY 358 RLPQLRVQPNLATMEYNLTRASKLFSLEQFIPTENTNFGNRL-VGISNRDAPTY--- 413
DB 320 LNALQEVGPTFSIENSIRKP--HLFDYLRGIEFHT-----RLRPYSGKDSFNWSG 370

QY 414 -----SNTIETLYGERTGSPYTKTIIRPFES---YKYSIVTDROSPPVSIQPH 459
DB 371 NYVETRPSIGSNDTITSPFYGDASKIEPIQKL--SFDGQKVYRTIANTDIAAFPDKI--Y 426

QY 460 FIINQIEL-YLNGSSNNLTLYSAGGSLSNYONTTFFQPRKDCNLVIDPGCS--PNFNN 516
DB 427 FGVTVDYFGYDDQKNET-----STQYDYSKRYNGYLGAQDSIDQLPEITDPELEXA 479

QY 517 YSHILSHFSLFTYSYVIGLQILDTGVLGWTHSSVDRYNAISDKIITMIPAIGNNLDT 576
DB 480 YSHQLNVAECFL-----MQDRRGITPFFTWTHRSVDFFNTIDAEKITQLPVVKAYALS 533

QY 577 NSKVIEBPGHGTGNLVYLO-----SQGRLEITCTPNSTQSYFIRLRVATNGAGNTLNI 632
DB 534 GASIIIEGPGFTGNNLTLKLESSNSIAKFVTLNSAALLQRYRIRYAST-----TNLR 587

QY 633 LITPGVIGIPPPQRLNNTFSGTNNYNNLQY--GDFGY--FQFPSTVTLPLNRNPIFNRA 688

Db 588 LFV-----QNSNNDFLVIYINKTNWIDGDLTYQTFDEATS-----NSNMGFSGDNTD 634

QY 689 -----VSNSTILIDKIEFIPI 704

Db 635 FIIGAESFVSNKIIYIDKIEFIPV 658

RESULT 6
I39811
parasporal crystal protein cry3Bb1 - Bacillus thuringiensis
N:Alternate names: parasporal crystal protein cryIIIB2
C:Species: Bacillus thuringiensis
C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 05-Oct-2004
C:Accession: I39811
R:Donovan, W.P.; Rupa, M.J.; Slaney, A.C.; Malvar, T.; Gawron-Burks, M.C.; Johnson, T.B.
Appl. Environ. Microbiol. 58, 3921-3927, 1992
A:Title: Characterization of two genes encoding Bacillus thuringiensis insecticidal crys
A:Reference number: I39811; MUID:93119147; PMID:1476436
A:Accession: I39811
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-652 <RES>
A:Cross-references: UNIPROT:Q06117; UNIPARC:UPI0000126C14; GB:M89794; NID:gl42729; PIDN:
C:Genetics:
A:Gene: cryIIIB2
C:Superfamily: Parasporal crystal protein

Query Match 16.9%; Score 654; DB 2; Length 652;
Best Local Similarity 27.8%; Pred. No. 2.5e-35;
Matches 206; Conservative 114; Mismatches 294; Indels 126; Gaps 27;

QY 1 MNQNDNNEYEIIDSTSPYFNNR--NSNDSRYPYTNPNQPLQNTYKWLWMCQ-QNT 57
DB 1 MNPNN-RSEHD-----TIKVTNSELTQNHQYPLADPNSTLELNKFEFLRMADNST 54

QY 58 QYDGNFETPASADTIA-AVSAGTIVSGTLAIGLGLTSISGPIGIIAIIISFGTLITVF 116
DB 55 EVLDN-----STVKDAGTGVGVQILGVG-----VFPAGALTFFYQSFLNTI 99

QY 117 WPAGEQDKTVWTOFIKMGEIFVDTPLTESIKQLKLTLEGPRQLQSYNTALDWRKLR 176
DB 100 WPS---DADPWKAFMAQVEVLIDKKIEYAKSKALAELOGLQNNFEDYVNALSNWKK--- 153

QY 177 LQAPGPPSALQQAALTAKIRENVHNDFIREIPGQLETYKTLPIYAQAANFHLNL 236
DB 154 -----TPLSRKRSQDRIRLEFSQAESHFNPSFAVSKFEVLFLPTVAQAANTHLL 208

QY 237 LQQAELADSEWADIHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKNLRDEP 296
DB 209 LKDAQVGEWYSGS-----SEDVAEFYHRQLKLT-QQYTDHCNVNWNVGLNLRGST 259

QY 297 NMKWSIFNDYRRVMTITVLDTISQFSLYDIKRYRDSIGGIEVKGIKNELTREIYTTTEINF 356
DB 260 YDAVVKENRRERMTLVLDLIVLPFYDRLYS-----KGVKTELTRDIETDPI-F 310

QY 357 DRLPQLRVQPNLATMEYNLTRASKLFSLEQFIPTY-----ENTNF--GNRLVGI 405
DB 311 SLNTLOBYGPTFSIENSIRKP--HLFDYLRGIEFHTRLQPGYFGKGSFNWSGNY---V 365

QY 406 SNRDAPYSNITITETLYGERTGSPYTKTIIRPFES---YKYSIVTDROSPP-----VSP 455
DB 366 ETRPSIGSNTITSPFYGDASKIEPIQKL--SFDGQKVYRTIANTDIAAFPDKI--Y 423

QY 456 IQPHFIINQIELYNGSSNNLTLYSAGGSLSNYONTTFFQPRKDCNLVIDPGCSNPN 515
DB 424 VD-----PSYDDQKNETSTQYDYSKRYNGYLGAQDSIDQLPEITD-----EPEK 470

QY 516 NYSHILSHFSLFTYSYVIGLQILDTGVLGWTHSSVDRYNAISDKIITMIPAIGNNL 575
DB 471 AYSHQLNVAECFL-----MQDRRGITPFFTWTHRSVDFFNTIDAEKITQLPVVKAYALS 524

QY 576 TNSKVIEBPGHGTGNLVYLO-----SQGRLEITCTPNSTQSYFIRLRVATNGAGNTL 631


```

Query Match      16.0%; Score 618; DB 2; Length 1160;
Best Local Similarity 26.6%; Pred No. 1.4e-32;
Matches 211; Conservative 116; Mismatches 293; Indels 174; Gaps 35;

Qy 5  NDNNEYIIDSHTSPYFPFNRNSDSRYFYTNPNPQLQNTYKWEIKNMCOG-NTQYGDNF 63
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4  NNQNEYIIDA-LSP--TSVSDNSIRYPLANDQTNTLQNMNKKDYLKMTSTNAELSRNP 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 64  ETPASADTIAVSAGTIVSGTLLAGIGGLTSGPIGII GAI IISFGTLITVFWPAGEQD 123
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61  GTFISAQD--AVGTGIDIVSTIISGLG-----IPVLGEVPSILGSLIGLLWPSNNEN 110
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 124  KTWMTQTKMGEIFVDTPLTESI KQLQTLGSGFRQILQSYNTALDDWRKLKRLQAGCLP 183
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 111  --VMQIFMNRVEELIDQKILDSVRSRAIDLANSRIAVRYQNALEDWRK-----N 159
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 184  PSSALQAAALTKIRFENVHNDFI REIPGFQLETKYKTLLLPIYAAANFHLNLQQGAEL 243
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 160  PHST--RSALVKERFGMAELRLTNMGSFSQNTYETPLPTTYQAASLHLLVNRDVQIY 217
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 244  ADEWNADIHPSQIBPNAGTSDDYIKLLKENIPKYSNYCANTYRTGKLNLEDENMKWSIF 303
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 218  GKEWG--YPO-----NDIDL FYKEQVSYTARYSDHCQVWYNAGLNKLRGTGAKQWVDY 268
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 304  NDYERYMTITVLDTISQPSLYDIKRYRDSIGGIEVKGIKNELTREIYTTEN----- 355
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 269  NRFRRENVMVLDLVALFPNYDARIY-----PLETNA-----ELTREIFTDPVSVYTGQSS 320
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 356  -----FRLPQLRVPQPNLATWEYNI TRASFKLPSFLEQFIYF-----ENTNF--GN 400
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 321  TLI SWYDMIP--AALPSPSTLENLRKPDF--FTLQEI RMWYTSFRQNGTIEYNYWGQ 376
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 401  RLVGISNRDAPTYSNTITELYGERTGSTPTTKTIRPFESYKVSIVTDQRQSPVSPQPHF 460
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 377  RL T-----LSYIYGS-----SFNKYSGVLGAGAEIDI PVGQNDIYR 411
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 461  I-----INOIEIYLNSSNNTLK-Y S-----AGG--SLSNYQNTFTFFQFPR 498
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```
Db 412 VVMTYIGRYTNSLLGVNPTFYF---SNTQKYSKPKQFAGGKIKTIDSGEELTY----- 463
QY 499 KDCNVLIDPGCSFNPNFNNYSHLSHFSLFTYSYVIGLQILDTGLVGTWTHSSVDYRNAI 558
Db 464 -----ENYQSYSHRVSYYITSPFIKSTGGTVLGVWP--IPGWTSSASRNNFI 508
QY 559 SDKIITMIPAIGKNNLDTSKVI-----EGPGHTGNNLVYLOSQ-----RLIEICEPNS 610
Db 509 YATKISQIPINKASR--TSGGAVNNFOEGL--YNGGPMVKLSGSGSQVINLRVATDAKAS 565
QY 611 QSYFIRLRVATNAGN--TLPNISLTIPGVIGIPQRLN--NTFSGTYNNLNQVGDGFGYFQ 667
Db 566 QRYRIRIRVADPAGKFTISSRSPENPATYSASIAVTNTMTWASLTYSTFYFAESG--- 622
QY 668 FSTVTPLNLRNIPFIRNADVS-----NSILIIDKIEFIPITSSMHQNRKQKLETIQ 721
Db 623 -----PINLGISGSSRTFDSITKEAGAANLYIDRIEIPVNTLF--EAEEDLDVAK 672
QY 722 TKINTFTNHTKTL 735
Db 673 KAVNGLFTNEKDAL 686

RESULT 9
S49247
parasporal crystal protein cry9Ca1 [validated] - Bacillus thuringiensis
N:Alternate names: parasporal crystal protein cryIh
C:Species: Bacillus thuringiensis
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 05-Oct-2004
C:Accession: A59350; S49247
R:Jambert, B.; Buyesse, L.; Decock, C.; Janssens, S.; Piens, C.; Saey, B.; Seurinck, ; V
Appl. Environ. Microbiol. 62, 80-86, 1996
A:Title: A Bacillus thuringiensis insecticidal crystal protein with a high activity agai
A:Reference number: A59350; MUID:96141404; PMID:857215
A:Accession: A59350
A:Molecule type: DNA
A:Residues: 1-1157 <LAW>
A:Cross-references: UNIPROT:Q45733; UNIPARC:UPI000002F5A5; EMBL:Z37527; NID:g547554; PID
A:Experimental source: serovar tolworthi
C:Comment: This parasporal crystal protein, active against corn borer and other ins: cte,
C:Superfamily: Parasporal crystal protein
C:Keywords: delta-endotoxin

Query Match 15.8%; Score 612; DB 1; Length 1157;
Best Local Similarity 26.6%; Pred. No. 3.6e-32;
Matches 209; Conservative 107; Mismatches 307; Indels 162; Gaps 30;

QY 1 MNQNDNNEYIIDSHTSYFPNRSNDRSPYTNPNPQNTNYKWLNMCGNTQYQ 60
Db 1 MNRNN-QNEYEIID--APHCGCPDSDVRYPLASDPNAALQNMKYKVLQMT--DEDYT 54
QY 61 DNF-ETPASADTIAVSACTIVSGTLLAGIGLTSIGPGIIGALLISFGTLITVFWPA 119
Db 55 DSYINPSLSISGRDAVQTALTVVGRIALG-----VPFSGQIVSVFYQFLNLTWEP 106
QY 120 GRQDKTWTQFTKMGEIFVDVDTPLTSTIKQLQTLGFRQILQSYNTALDDW-----R 172
Db 107 --NDTAWEAFNRQVEELVNOQITFARNQALRLQGLGDSFNRYORSQNWLADNRDTR 164
QY 173 KLRKLAQAPGLPSSALQQAALTKIRFENVHNDFIREIPGFQLETYKTLILLPIYAQAANF 232
Db 165 NLSVVRA-----QFTALDL-----DFVNAIPLFAVNGQOVPVLLSYAQAANL 206
QY 233 HNLILQOQALADENADIHPSQIEPNAGTSDDYKLLKENIPKYSNCANTYRTGLKNL 292
Db 207 HLLLLKDALSLFGEHGF-----TQGEISTYVDRLQELTAKYNYCETWYNTGLDRL 257
QY 293 RDEPNMKISIFNDYRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKGKINELTRIYTT 352
Db 258 RGTNTESMLRYHQFRREMTFLVLDVVALFPFYVDRLY-----PTGSNPQLTREYVTD 309
QY 353 EINFDRLPQLRVQPNLATMEYN---LTRASFK---LFSFLEQFIF---YTENTNFGNR 401
```

```
Db 310 PIVFNPPANVGLCRWGTNPYNTFSELENAFIRPPHLPDRLNSLTISNRRFPVSSNFM 369
QY 402 LVGISNRDAPTYGNTITETILYGBRTGSPPTTKTRP-----FESYKVSIVTD 447
Db 370 WSGHTLRSYLNDSAVQEDSYGLIT--TTRATINPGVDGTRNRIESTAVDFPSALIGI--- 424
QY 448 RQSPVPSPIQPHFIINOIELYLNQSSNNTLYKASAGSLSNYQNTFFQFPKPKDCNVL 507
Db 425 -----YGVNRASFVPGGLEFNGTTPSPANGCGRDLY-----DTNDEL 460
QY 508 PGCSNPNFNNYSHLSHFSLFTYSYVIGLQILDTGLV-----WTHSSVDYRNAI 558
Db 461 P--DESGSTHRLSLSHVTFFSF-----QTNQAGSIANAGSVPTVYVWTRDRDLNNTI 510
QY 559 SDKIITMIPAIGKNNLDTSKVI-----EGPGHTGNNLVYLOSQ-----GRLEITCETPNSTQSYFI 615
Db 511 TPNRIITQLPLVKASAPVSGTTLVKGFGTGGILRTTNGTFTGLRTVTVNSP-ITQQYRL 569
QY 616 RLRYATNGAGNTLPNISL-TIPGVIGIPQRLNNTFSGTYNNLNQVGDGFGYFQFPSTVTL 674
Db 570 RVRFASGTG-----NFSIRLVRCGVSIGDVRGSGTMN--RQQLTYESFFTFREF--TTTG 619
QY 675 PLNRNIPFIRNADVNSI-----LIIDKIEFIPITSSMHQNRKQKLETIQTKI 724
Db 620 PFNP--PFTTQAEILTVNAEGVSTGGEYVIDRIEIVPVNPA---REABEDLEAKKAV 674
QY 725 NTFET 729
Db 675 ASLFT 679

RESULT 10
A27323
parasporal crystal protein cry3Aal - Bacillus thuringiensis
N:Alternate names: coleopteran-specific insect control protein; cryetal protein cryC; del
C:Species: Bacillus thuringiensis
C>Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 05-Oct-2004
C:Accession: A27323; A26853; A29987; A28407; S60781; I39813
R:Herrnstadt, C.; Gilroy, T.E.; Sobieski, D.A.; Bennett, B.D.; Gaertner, F.H.
Gene 57, 37-46, 1987
A:Title: Nucleotide sequence and deduced amino acid sequence of a coleopteran-active del
A:Reference number: A27323; MUID:88112860; PMID:2828180
A:Accession: A27323
A:Molecule type: DNA
A:Residues: 1-652 <HER>
A:Cross-references: UNIPROT:Q9S6N9; UNIPARC:UPI000000AEEEDC; GB:M22472; NID:gl42733; PIDN:/
A:Experimental source: strain San Diego
R:Roette, H.; Seurinck, J.; Van Houtven, A.; Vaeck, M.
Nucleic Acids Res. 15, 7183, 1987
A:Title: Nucleotide sequence of a gene encoding an insecticidal protein of Bacillus thuri
A:Reference number: A26853; MUID:88015559; PMID:3658680
A:Accession: A26853
A:Molecule type: DNA
A:Residues: 9-652 <HOE>
A:Cross-references: UNIPARC:UPI000002B8CD; GB:Y00420; NID:g40252; PIDN:CAA68482.1; PID:g
A:Experimental source: var. tenebrionis
R:McPherson, S.A.; Perlak, F.O.; Fuchs, R.L.; Marrone, P.G.; Lavrik, P.B.; Fischhoff, D.J.
Bio/Technology 6, 61-66, 1988
A:Title: Characterization of the coleopteran-specific protein gene of Bacillus thuringier
A:Reference number: A29987
A:Accession: A29987
A:Molecule type: DNA
A:Residues: 9-652 <MCP>
A:Cross-references: UNIPARC:UPI000002B8CD
A:Experimental source: var. tenebrionis
R:Sekar, V.; Thompson, D.V.; Maroney, M.J.; Bookland, R.G.; Adang, M.J.
Proc. Natl. Acad. Sci. U.S.A. 84, 7036-7040, 1987
A:Title: Molecular cloning and characterization of the insecticidal crystal protein gene
A:Reference number: A28407
A:Accession: A28407
A:Molecule type: DNA
A:Residues: 9-652 <SEK>
```

A;Cross-references: UNIPARC:UPI000002B8CD
A;Experimental source: var. Tenebrionis
R;Adams, L.F.; Mathewes, S.; O'Hara, P.; Petersen, A.; Guertler, H.
Mol. Microbiol. 14, 381-389, 1994
A;Title: Elucidation of the mechanism of CryIIIA overproduction in a mutagenized strain
A;Reference number: S60781; MUID:95131759; PMID:7830581
A;Accession: S60781
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 9-652 <ADA>
A;Cross-references: UNIPARC:UPI000002B8CD; EMBL:U10985; NID:9506182; PIDN:AAQ432866.1; PT
A;Experimental source: var. tenebrionis strain NB176, a mutant of strain NB125
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1994
R;Wu, S.J.; Dean, D.H.
J. Mol. Biol. 255, 628-640, 1996
A;Title: Functional significance of loops in the receptor binding domain of Bacillus thuringiensis
A;Reference number: S62317; MUID:96163559; PMID:8568902
A;Contents: annotation
R;Donovan, W.P.; Gonzalez, J.M.
Mol. Gen. Genet. 214, 365-372, 1988
A;Title: Isolation and Characterization of EG2158, a new strain of Bacillus thuringiensis
A;Reference number: I39812; MUID:89112139; PMID:3146015
A;Accession: I39812
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 9-652 <RES>
A;Cross-references: UNIPARC:UPI000002B8CD; GB:M37207; NID:g142735; PIDN:AAA50255.1; PID:
A;Experimental source: strain EG2158
R;Teixeira De Souza, M.; Lecadet, M.M.; Lereclus, D.
J. Bacteriol. 175, 2952-2960, 1993
A;Title: Full expression of the cryIIIA toxin gene of Bacillus thuringiensis requires a
A;Reference number: I39813; MUID:93259939; PMID:8491716
A;Accession: I39813
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 9-58 <RE2>
A;Cross-references: UNIPARC:UPI000008007B; GB:L03393; NID:g304150; PIDN:AAA22350.1; PID:
C;Genetics:
A;Gene: cryIIIA
C;Superfamily: Parasporal crystal protein
C;Keywords: delta-endotoxin

Query Match 15.3%; Score 592.5; DB 2; Length 652;
Best Local Similarity 26.4%; Pred. No. 2.9e-31;
Matches 196; Conservative 101; Mismatches 307; Indels 139; Gaps 27;

QY 1 MNQNDNNEVEIIDSHTSPYFPNNSDSDRYPTNNPQNTQNTYKWLNMCGNTQYG 60
DB 9 MNPNN-RSEHDTTKTENNEVP---TNHVQYPLAETPNPTLEDNLNKEFLRMWTDN--- 60

QY 61 DNFETPASADTIAAASAGTIVSGTLLAGIGGLTSISGPIIGIAGIIISFGTLITVFWPAG 120
DB 61 -NTEALDSSTTKDVIQKGSIVGDLGVVG-----FPFG--GALVSFYTNFLTITWPE 111

QY 121 EODKTVWTFQIKMGEIFVDTPTLSTESIKQLKLOTLEGFRQILQSYNTALDDWRKLRLOAP 180
DB 112 DP-----WKAFMEQVEALMDQKIADYAKNRKALAEQLQNNVEDVYSSWQK----- 160

QY 181 GLPPSSALQQAALTALKIRFNVDNFIPIPGFQLETYKTLILLPIYAAQANFHLNLLQQ 240
DB 161 -NPVSSRNPHSQRIELFSQAESHFRNSMPSFASIGYEVLFLTYYAQANFHLNLLKDA 219

QY 241 AELADEW---NADIHSQIEPNAGTSDDYVKKLKENIPKYSNYCANTYRTGLKNLRDEPN 297
DB 220 QYGEWGEYKEDI-----AEFYKRLKLTQETDHCVRKYNVGLDKLRGSSY 267

QY 298 MKWSIFNDYRRYWTITVLDTISQFSLYDIKRYRDSIGGIEVKIGKKNELREIYVTEI--- 354
DB 268 ESWVFNRYRRMTLVLDLIAFLPYDYRLY-----PKEVKTELTRDVLTDPIGV 319

QY 355 -----NFDRLPQLRQVQPNLATMEYNLATRASFKLSFLQEIFYTE----- 394
DB 320 NNLRGYGTTFNSNIENIRKPH-----LPDYLRHQETFRFQPGY 359

RESULT 11

S00873

N;Alternate names: parasporal crystal protein cryA4
C;Species: Bacillus thuringiensis subsp. thuringiensis
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 05-Oct-2004
C;Accession: S00873

R;Brizzard, B.L.; Whiteley, H.R.

Nucleic Acids Res. 16, 2723-2724, 1988

A;Title: Nucleotide sequence of an additional crystal protein gene cloned from Bacillus t
A;Reference number: S00873; MUID:88203216; PMID:3362680
A;Molecule type: DNA

A;Accession: S00873

A;Residues: 1-1228 <BRI>

A;Cross-references: UNIPROT:P05517; UNIPARC:UPI0000126BDC; EMBL:X06711; NID:g40264; PIDN:

C;Genetics:

A;Gene: cryA4

A;Start codon: TTG

C;Superfamily: Parasporal crystal protein

C;Keywords: delta-endotoxin

Query Match 14.9%; Score 576; DB 2; Length 1228;

Best Local Similarity 25.6%; Pred. No. 9.4e-30;

Matches 202; Conservative 107; Mismatches 285; Indels 196; Gaps 32;

Db 51 VEPFVSASTTQ-----TGIGAGKILGTGVFPAGQVASLSPILGELPKG 97

Qy 121 EQDKTVMTQFIKMGEEIFVDTPLETSIKQLKLTLEGFRQILQSYNTALDDW---RKLKRL 177

Db 98 ---KNQWEIEMEHVEEIIINQKISTYARNKALTDLKGLDALAVYHDSLESWGNRNTRA 154

Qy 178 QAPGLPSSALQQAALTKIRFENVHNDFTREIPGQLETYKTLTLLPIYAAANFHNLL 237

Db 155 R-----SVVKSQYIAELM-----FVKLPFAVSGEEVPLPIYAAANLHLLL 200

Qy 238 QCGAELADEWADLHPSQIEPNAGTSDDYKLLKENIPKYSNYCANYRGLKLNLRDEPN 297

Db 201 RDAISIFKENG--LSSSEI-----STFNQRQVERAGDYSYHCWKWYSGLNLRGTNA 251

Qy 298 MKWSIFNDYRRYMTITVLDTISQSLYDIKRYRDSIGGIEVKGIKN--ELTREIYTTTEIN 355

Db 252 ESWRVYNQPRDMLMVLDLVALPFSYDTQWY-----PIKTTAQLTREYVT--- 297

Qy 356 FDLPLQRLVQPNL-ATWEYNLTRASPK-----LFSFLEQFIPT-----ENTNF 398

Db 298 -DAIGTVHPHPSFTSTTWYNNAPSPSAIEAAVVRNPHLLDFLEQVFIYSLLSWSNTQY 356

Qy 399 GNLVIGISNRDAPYSNTITETLVGERTGSPPTTKTIRPEES---YKVSIVTDQ---SPP 452

Db 357 MN-MWGGHKLEFRTIGTILNISTOGSTNTSINPVTL-PFTSRDVRVYTESLAGMLFLTQP 414

Qy 453 VSPIQP-----HFIINOI---ELYINGSSNNLTKYSAGGSLSNYQNTFFQPRKKDCNL 504

Db 415 VNGVPRVDHFWKFWTHPIASDNFYPGYA-----GIGTQLQDSEN-----EL 456

Qy 505 VIDPGSPFNFNYSHIULSHFSLFTYSYVIGLQILDTGVLGWTHSSVDRYNAISKIIT 564

Db 457 PPEATGQPNYESHRLSHLGLISASHVKAL-----VYSWTHRSADRTNIEPNSIT 508

Qy 565 MIPAIGKNLDTNSKVTEGPGHTGKNLVYLQSQ-----RLEITCEPNSTOSYFIRLY 619

Db 509 QIPLVKAFNLSSGAAVRGPGFTGGDILRNTNTTFGDIKVI---NPPFAQRYRVIRY 565

Qy 620 AT-----NGAGNTLPIISLITPGVIGIPQRLNNTFSGTNYNNLQVGFQGYQFP 669

Db 566 ASITDLQFHTSINGKAINQGNFSATM-----NRGEDLDYKTFRTVGTFT 608

Qy 670 STVTPLPLNRNIPFNRAVNSNLI-----IDKIEFIPITSSMHQNRKQKL 717

Db 609 T-----PFSF--LDVQSTFTIGAMNFSNGEYVIDRIEFVPEVVEYEAED---F 653

Qy 718 ETIQTKINTEFTN 730

Db 654 EKAQEKVLTALTS 666

RESULT 14

S25383

Parasporal crystal protein cryIIa1 - Bacillus thuringiensis

N;Alternate names: delta-endotoxin; parasporal crystal protein cry7

C;Species: Bacillus thuringiensis

C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 05-Oct-2004

C;Accession: S25383

R;Tailor, R.; Tippet, J.; Gibb, G.; Pells, S.; Pike, D.; Jordan, L.; Ely, S.

Mol.: Microbio1, 6, 1211-1217, 1992

A;Title: Identification and characterization of a novel Bacillus thuringiensis delta-endotoxin

A;Reference number: S25383; MUID:92269582; PMID:1588820

A;Accession: S25383

A;Molecule type: DNA

A;Residues: 1-719 <TAI>

A;Cross-references: UNIPROT:Q45752; UNIPARC:UPI0000002DB73; EMBL:X62821; NID:940289; PDB:1A01

C;Genetics:

C;Superfamily: Parasporal crystal protein

C;Keywords: delta-endotoxin

Query Match 14.6%; Score 565.5; DB 2; Length 719;

Best Local Similarity 26.2%; Pred. No. 2.1e-29;

Matches 208; Conservative 106; Mismatches 284; Indels 195; Gaps 35;

```
Qy 3 QNNDNNYEYIIDSHTSYPFNRNSDSRYFVTNPNQPLQNTNYKEWLNWCOGNTOYGDN 62
Db 4 KNQDKHQ-----SFSSNAKVDKISTDS---LKNETDIELQNIHEDCLKM-----SEY-EN 50
Qy 63 FETPASADTTAAVSAGTIVSGTLLAGTGTSTSGPIGI--IGAIISFPGTLITVFWPAG 120
Db 51 VEPFVSASTIQ-----TGIGIAGKILGTGVPFAGQVASLYSFILGELWPKG 97
Qy 121 EODKTVMVTFKMGIEIFVDTPLTESIKNQKLQTLLEGPRQILQSNVNTALDWM---RKLKRL 177
Db 98 ---KNQWEIEMEHVEEIIINOKISTYARNKALTDLKGLGDALA VYHDSLESVWGNRNNTRA 154
Qy 178 QAPGLPSSALQQAALTKLIRFENVHNDFFIREIPGFOLEYTKYTLTLLPIYAAANFHLNLL 237
Db 155 R-----SVKSYIALELM-----FVQKLPSFAVSGEEVPLPIYAAANLHLLLL 200
Qy 238 QQGALADENADIHPSQIEPNAGTSDDYKYLKENIPKYSNVCANTYRTGLKNLRDEPN 297
Db 201 RDAISIFGKEWG--LSSSEI-----STFYNRQVERAGDYSYHCVKWYSTGLANLRGTNA 251
Qy 298 MKYSIENDERYMTITVLDTISQFSLYDKRYRDSIGGIEVKIGN--ELTREIYTYTEIN 355
Db 252 ESWRVYNQFRDWTLMVLDLVALFPSTDTOMY-----PIKTAQTREYVT----- 297
Qy 356 FDLRLQRLRVQPNL-ATMEYNLTRASPK-----LPSFLEQFIYFT-----ENTNF 398
Db 298 -DAIGTVHPHPSFTSTWYNNNAPSFAIEAAVVRNPHLLDFLEQVTIYSLLSRWSTQY 356
Qy 399 GNRLVGINSRDAPYSNTITETLYGBERTGSGPTTKTIRPPES-----YKVSIVTDRO---SPP 452
Db 357 MN-MWGGHKLEFRITIGTGLNISTQGSNTSINPVTI--PFTSRDVRITESLAGNLFLTQP 414
Qy 453 VSPIQP-----HFLINQI---ELYLNGSSNNTLYKSAGSLSNVQNTTFFQFPRKKDCNL 504
Db 415 VNGVRVDFHWKVFVTHPIASDNFYYPGYA-----GIGTQLQSEN-----EL 456
Qy 505 VIDPGCSPNFNNTSHILSHPSLFETYXVIGLQQLDGLVGLWTHGSDVRYNIAISDKIIT 564
Db 457 PPEATGQPNYESXSHRLSHLGLISASHVKAL-----VYSWTHRSADRTNIEPSIT 508
Qy 565 MIPAIGNNLDNTNSKVIEGSGHTGNGNLVYLSQSG-----RLEITCTPNSTQSYFIRLRY 619
Db 509 QIPLVKAFLNLSGAAVVRGPGFTGGDILRRTNTGTGDIRVNI---NPPPAQRYRVRIY 565
Qy 620 AT-----NCAGNTLPNISLITPGVIGIPQRLNNTFSGTNYNNLOYGDFGFGVQFP 669
Db 566 ASTTDLOFHTSINGKAINQGNFSATM-----NRGEDLYKFTFTVGVGT 608
Qy 670 STVTPLNRNIPFIFNRADYSNGLI-----IDKIEFIPITSSMHQNRQKL 717
Db 609 T-----PSPF--LDVQSFTTGAMWFGSSGNEVYIDRIEVPVEVTEAEYD---F 653
Qy 718 ETIQTKINTPFTN 730
Db 654 EKAQEKVTLFTS 666
```

RESULT 15
JH0261
parasporal crystal protein cry3Cal - *Bacillus thuringiensis* subsp. *kurstaki* (strain
N; Alternate names: parasporal crystal protein cryIIID
C; Species: *Bacillus thuringiensis* subsp. *kurstaki*
C; Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 05-Oct-2004
C; Accession: JH0261; S18944
R; Lambert, B.; Theuniss, W.; Aguda, R.; Van Audenhove, K.; Decock, C.; Janssens,
Gene 110, 131-132, 1992
A; Title: Nucleotide sequence of gene cryIIID encoding a novel coleopteran-activ
A; Reference number: JH0261; MUID:92184108; PMID:1544571
A; Accession: JH0261
A; Molecule type: DNA
A; Residues: 1-649 <LAM>

A;Cross-references: UNIPROT:Q45744; UNIPARC:UPI0000126C15; EMBL:X59737; NID:g40287; PIDN:
C;Genetics:
C;Gene: cryIIID
C;Superfamily: Parasporal crystal protein

Query Match	14.6%	Score 564;	DB 1;	Length 649;
Best Local Similarity	26.2%	Pred. No. 2.2e-23;		
Matches 203;	Conservative	98;	Mismatches 275;	Indels 200; Gaps 31;
Qy	1	MNQNDNNEVEIDSHTSPIYFPNRSNDSRVPYTNPNQPLQNTNYKEWLNMCQNTQYG	60	
Db	1	MNPNN-RSEHDTIKATENNEVSN--NHAQIPLADTPT--DEELNYKEFURRTDN----	50	
Qy	61	DNEETASADTIAAASAGTIVSGTLLAGIGGLTISGPIGIGIAIIISFGTLITVPWPAG	120	
Db	51	NVEALDSSTTKDAIKQGISIIGLLGVG-----FPYG--GALVSFYTNLLNTIWPGE	101	
Qy	121	EQKTYWTFPIKGEIIFVDPLTESIKQLKLTLEGFRQILQSNYNTALDWRKUKRLQAP	180	
Db	102	DPLKA-----FMQOQEALIDQKIADYAKDKATAELQGLKNVFKOYVSALDWDK-----	150	
Qy	181	GLPPSSALQQAALTL-----KIR--FENVHNDPIREIPGOLEUTKYTKLLLPYAAQAN	231	
Db	151	-----TPLTRDGRSGRIREIFPSQAESHFRRSMPSFAVSGEVLFLPYAAQAN	200	
Qy	232	FHLNLQOGAELADEWNADIHPSQIBENAGTSD--DYKLLKENIPKYSNYCANTYRTGL	289	
Db	201	THLLLLKDAQIYGTDW-----GYSTDLLNEFHTKQKOLTIETYNHCAKWYKAGL	249	
Qy	290	KNLRDEPNKWSIPNDYRRYNTITVLDTISQFSLYDIKRYRDSICGIEVGKIKELTREI	349	
Db	250	DKURGTYEEWVKFNRYRREMTLPLDLITLFPDYVRTY-----TKGVKTELTRDV	301	
Qy	350	YTTEI-----NEDRLPOLKRVQPNLATMEYNLTRASFKLPSLEQEIFY	392	
Db	302	LTDPIAVNMNGVITTFSENIVIRPHL-----FDY--LHAIQFHSRLQPGVYG	350	
Qy	393	TENTNF--GNRLVGISNRDAPYNTTITETLYGBRTSGPTTKTIR--PFBSYKVSIVTDR	448	
Db	351	TDSPNYWSGNY--VSTRSSIGSDEIIRSPFYGNK-----STLDVQNLEFNGEKVFRAVAN	403	
Qy	449	QSPVPSPi-----OPHFIINOELIYLCNGSSNNTLKYSAGGSLSNYQNTTFQF-----PRK	499	
Db	404	GNLAVPWGTGTHKHSQVTKVQSPQVNDKRDVTRQTYDSKRVNGGIVFDSIDQLPPIIT	463	
Qy	500	KOCNLVIDPGCSPNFNYSHILSFSLFTYSYVIGLQQLIDTGC-----VLGWTHTSSVDR	554	
Db	464	TDESLE-----KAYSHQNLNVRCF-----LLQGGRIIPVFTWTHKSVDFF	503	
Qy	555	YNAISDKLIITWIPAKGNLNDTNSKVIPEGPCHTGCNLYLOSQGRLEI--TCETPNSTQS-	612	
Db	504	YNTLDSKTIQIPFVKAFILVNSTSVVAGPGFTGGDIIKCTNGSGLTYLYVPAPDLTYSK	563	
Qy	613	-YFRLRYA-----TNAGNTLPNIULTIPGVIGIPQRLNN	648	
Db	564	TYKIRIYASTSQVRFGIDLGSYTHSISYEDKTDKXNLTLYSNFNLSSV-SRPIE-----	618	
Qy	649	TFSGTNYNNLQYGDYFQFPFSTVTLPLNRNPIPIFNRADVNSILIIDKIEFTPI	704	
Db	619	-ISGNGKIGYSVGGIG-----SGDEVYIDKIEFTPM	648	

Search completed: December 4, 2005, 13:12:34
Job time : 45 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2005, 12:51:52 ; Search time 230 Seconds
(without alignments)

2254.623 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869

Sequence: 1 MQQNNDNVEIIDSHTSPY.....KLETIQTKINTFFNHTKTL 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1060.5	27.4	1180	1	P16480 bacillus th
2	1060.5	27.4	1180	2	Q7AL67 BACTI
3	892.5	23.1	675	1	C10AA BACTI
4	892.5	23.1	675	2	Q8KNV2 BACTI
5	862.5	22.3	1128	2	Q9FDC0 BACTF
6	849	21.9	1109	1	C28AA BACTF
7	847	21.9	650	2	Q8VNX2 BACTV
8	803.5	20.8	683	2	Q75VA2 BACTE
9	784	20.3	688	2	Q8VNX1 BACTV
10	775.5	20.0	1169	1	CR8BA BACUK
11	758.5	19.6	1136	1	CR4BA BACTI
12	758.5	19.6	1136	2	Q7AL72 BACTI
13	734	19.0	688	2	Q5W7N9 BACTU
14	709	18.3	682	1	C19BA BACUH
15	705	18.2	829	2	Q6BE06 BACTU
16	677	17.5	825	2	Q6BE09 BACTU
17	676	17.5	659	1	CR3BA BACTO
18	672.5	17.4	648	1	C19AA BACTU
19	663	17.1	1169	2	Q56B08 BACTU
20	662	17.1	660	2	CR8OU6 BACTA
21	660.5	17.1	1138	1	CR7AB BACUK
22	654	16.9	652	1	CR3BB BACTU
23	649	16.8	686	2	Q75Q05 BACTE
24	641.5	16.6	1138	1	CR7AA BACTU
25	639.5	16.5	1280	2	Q8VUK9 BACTU
26	633	16.4	1144	2	Q8KZL7 BACTG
27	632.5	16.3	1138	1	CR7AB BACUA
28	630	16.3	1236	2	Q939T3 BACTU
29	619	16.0	1231	2	Q8KNY2 BACTU
30	618.5	16.0	1169	1	CR9DA BACTP
31	618	16.0	1160	1	CR8CA BACTP

32	618	16.0	1160	2	Q6R2R6 BACTU	Q6r2r6 bacillus th
33	615.5	15.9	1157	1	CR8AA BACUK	Q45704 bacillus th
34	615	15.9	1231	1	CR1BD BACTZ	Q9zaz5 bacillus th
35	612.5	15.8	675	2	Q6BCH5 BACTU	Q6bch5 bacillus th
36	612	15.8	1157	1	CR9CA BACTO	Q45733 bacillus th
37	604	15.6	1163	2	Q5XLA8 BACTP	Q5xla8 bacillus th
38	599	15.5	826	1	C27AA BACUH	Q9a597 bacillus th
39	597	15.4	1340	2	Q589X2 PARPP	Q589x2 paenibacill
40	595	15.4	1344	2	Q76SX7_9BACL	Q76sx7 paenibacill
41	594.5	15.4	652	2	Q6FXN8 BACTU	Q6pxn8 bacillus th
42	592.5	15.3	644	1	CR3AA BACTM	P0a381 bacillus th
43	592.5	15.3	644	1	CR3AA BACTM	P0a380 bacillus th
44	592.5	15.3	644	1	CR3AA BACTT	P0a379 bacillus th
45	592.5	15.3	652	2	Q9S6N9 BACTU	Q9s6n9 bacillus th

ALIGNMENTS

RESULT 1
ID CR4AA_BACTI STANDARD; PRT; 1180 AA.
AC P16480;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Pesticidal crystal protein cry4Aa (insecticidal delta-endotoxin
DE CryIVA(a)) (Crystalline entomocidal protoxin) (135 kDa crystal
DE protein).
GN Name=cry4Aa; Synonyms=cryIVA(a), isrH4;
OS Bacillus thuringiensis subsp. israelensis.
OG Plasmid 72 Kb.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=1430;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sen K., Honda G., Koyama N., Nishida M., Neki A., Sakai H., Himeno M.,
RA Komano T.;
RT "Cloning and nucleotide sequences of the two 130 kDa insecticidal
RT protein genes of Bacillus thuringiensis var. israelensis.";
RL Agric. Biol. Chem. 52:873-878 (1988).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA MEDLINE=88015571; PubMed=2821500;
RX Ward E.S., Ellar D.J.;
RT "Nucleotide sequence of a Bacillus thuringiensis var. israelensis gene
RT encoding a 130 kDa delta-endotoxin.";
RL Nucleic Acids Res. 15:7195-7195 (1987).
RN [3]
RP MUTAGENESIS STUDIES.
RX MEDLINE=94307434; PubMed=7913448; DOI=10.1016/0014-5793(94)00604-0;
RX Nishimoto T., Yoshiaue H., Ihara K., Sakai H., Komano T.;
RT "Functional analysis of block 5, one of the highly conserved amino
RT acid sequences in the 130-kDa CryIVA protein produced by Bacillus
RT thuringiensis subsp. israelensis.";
RL FEBS Lett. 348:249-254 (1994).
CC -!- FUNCTION: Promotes colloidotonic lysis by binding to the midgut
CC epithelial cells of insects.
CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -!- MISCELLANEOUS: Diverse amino acid mutations in sequence block 667-
CC 676 have no direct effect on the insecticidal activity but alter
CC the structural stability of the toxin protein molecule.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not

```
CC removed.
DR EMBL; D00248; BAA00179.1; -; Genomic DNA.
DR EMBL; Y00423; CAA68485.1; -; Genomic DNA.
DR PIR; A26858; A26858.
DR PIR; I39870; I39870.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin.C.
DR InterPro; IPR005639; endotoxin.N.
DR Pfam; PF03944; Endotoxin.C; 1.
DR Pfam; PF00555; Endotoxin.M; 1.
DR Pfam; PF03945; Endotoxin.N; 1.
DR Pfam; PF03945; Endotoxin.N; 1.
KW Plasmid; Sporulation; Toxin.
FT CONFLICT 306 306 L -> V (in Ref. 2).
FT CONFLICT 1052 1052 T -> I (in Ref. 2).
FT CONFLICT 1109 1109 R -> G (in Ref. 2).
FT CONFLICT 1127 1127 C -> W (in Ref. 2).
SQ SEQUENCE 1180 AA; 134539 MW; 6FB5B6979DACAD3B CRC64;

Query Match 27.4%; Score 1060.5; DB 1; Length 1180;
Best Local Similarity 35.1%; Pred. No. 1.1e-59;
Matches 271; Conservative 118; Mismatches 278; Indels 105; Gaps 24;

QY 1 MNQNDNNEVEIIDSHTSYFPNRRNSDRPYTNPNQPLQNTNYKEWLNMCQNTQYG 60
DB 1 MNPYQNKNEYETLNASQKKL--NISNTRYPIENSPKQLQSTNYKDWLNMCQNTQYG 58
QY 61 DNFETPASADTIAAAGTIVSGTLGAGIGGLTSISGPIGIIAIIISFGTLITVFWPAG 120
DB 59 GDFETFIDS--GELSAITIVGTVLTGFGFT----PLGL--ALIGFGLTLPVLPFAQ 108
QY 121 EQDKTWTQFIKMGEIFVDTPLTESIKQLKQLEGRQILOSYNTALDWRKLRLOAP 180
DB 109 DQSNLT--WSDFITQTKNIKKEIASTVISNANKILNRSFNVISTYHNHLKTWE----NNP 162
QY 181 GLPPSSALQQAALTALKIRENVHNDIFREIP--GFQLEYTKTLLPIYAAQANFHLNLQ 238
DB 163 NPQNTQDVRTQIQLVHYHFQNVIPELVNSCPPNPSDCDYNNILVLSYAAQANLHLTVLN 222
QY 239 QGAELADEWNADIHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKNLRDEP-- 296
DB 223 QAVKEAYLKNRRQFDYLEP-LPTAIDYVPVLTAKIEDYNYCVTVTKYKGLNLKTTTDS 281
QY 297 ----NMKWSIFNDYRYMTITVLDTISQPSLYDIKRYRDSIGGIEVKIGKNELTREIYTT 352
DB 282 NLDGNINWNTYRTKMTTAVLDLVALFPNDVGKY-----PIGVQSELTRREIYQV 333
QY 353 EINFDRLPOLRVQPNLATMEYNLTRASFLEQFIY-----TENTNFGNR----- 401
DB 334 -LNFESPYKYY--DFQYQEDSLTRRP-HLFTWLDLSLNFYEKAQTTPNPFTHYNNFHY 389
QY 402 -LVGISNRDAPYSNTITETLYGERGTSPTTKTIRPFESYKYSIVTDROSPPVSPQPHF 460
DB 390 TLDNISQSSVFGNHNVTDKL--KSLGLATNIYI-----F 422
QY 461 IINQIEL---YLNQSSNNT-----LKYSAGGSLSNYQ-NTTFFQPRKKDC 502
DB 423 LLNVISLDNKNYLDNYSKMPFFITNGTRLEKELTAGSGQITVDYVKNKIFGLPKRR 482
QY 503 NLVIDPGCSNPNFNYSKSHLSHSLFTYSYVIGIQLQILDGTGLVGTWTHSSVDYRNAISDKI 562
DB 483 ENQGNPTLPFTYDNYSHILSFIKSLSPATYKTQVY-----TFAWTHSSVDPKNTIYTHL 537
QY 563 ITMIPAIKGNLDTNSKVIETGPGHTGCGNLYVLOSQGRLEITCETPNSTQSYFIRLYATN 622
DB 538 TFOIIPAVKANSIGTAKSVQVGGPGHTGGDLI--DFKDHFKITQHSFNQFSYFIRIYASN 595
QY 623 GAGNTLPNTSLTPGVIGIPQRLNNTFSGTNNLYQVDFGVGFQFPSTVTLPLNRPNIP 682
DB 596 GSANTRAVNLISIPGVAEL--GHALNPTFGSDTNNLYKYKDFQYLEFNEVKFAPNQISL 654
QY 683 IFNRADV--SNSLIIDIKBIFIPITSSMQNREKQKLETIQTKINTFTNHTK 733
DB 683 IFNRADV--SNSLIIDIKBIFIPITSSMQNREKQKLETIQTKINTFTNHTK 733
```

```
DB 655 VFNRSVDYVNTTTLVDIKIEFLPITRSIREDEKQKLETVQOQINTFYANPIK 706

RESULT 2
Q7AL67_BACTI
ID Q7AL67_BACTI PRELIMINARY; PRT; 1180 AA.
AC Q7AL67_BACTI
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Pestidial crystal protein cry4AA.
GN Name-cry4AA; Synonyms-cryIVA(A), isrH4, p8t110;
OS Bacillus thuringiensis (subsp. israelensis).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=1430;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22235415; PubMed=12324359;
RX DOI=10.1128/AEM.68.10.5082-5095.2002;
RA Berry C., O'Neil S., Ben-Dov E., Jones A.F., Murphy L., Quail M.A.,
RA Harris D., Zaritsky A., Parkhill J.;
RT "Complete sequence and organisation of pBtoxis, the toxin-coding
RT plasmid of Bacillus thuringiensis subsp. israelensis.";
RL Appl. Environ. Microbiol. 68:5082-5095(2002).
DR EMBL; AL731825; CAD30148.1; -; Genomic DNA.
SQ SEQUENCE 1180 AA; 134538 MW; 6FB5B6979DACAD3B CRC64;

Query Match 27.4%; Score 1060.5; DB 2; Length 1180;
Best Local Similarity 35.1%; Pred. No. 1.1e-59;
Matches 271; Conservative 118; Mismatches 278; Indels 105; Gaps 24;

QY 1 MNQNDNNEVEIIDSHTSYFPNRRNSDRPYTNPNQPLQNTNYKEWLNMCQNTQYG 60
DB 1 MNPYQNKNEYETLNASQKKL--NISNTRYPIENSPKQLQSTNYKDWLNMCQNTQYG 58
QY 61 DNFETPASADTIAAAGTIVSGTLGAGIGGLTSISGPIGIIAIIISFGTLITVFWPAG 120
DB 59 GDFETFIDS--GELSAITIVGTVLTGFGFT----PLGL--ALIGFGLTLPVLPFAQ 108
QY 121 EQDKTWTQFIKMGEIFVDTPLTESIKQLKQLEGRQILOSYNTALDWRKLRLOAP 180
DB 109 DQSNLT--WSDFITQTKNIKKEIASTVISNANKILNRSFNVISTYHNHLKTWE----NNP 162
QY 181 GLPPSSALQQAALTALKIRENVHNDIFREIP--GFQLEYTKTLLPIYAAQANFHLNLQ 238
DB 163 NPQNTQDVRTQIQLVHYHFQNVIPELVNSCPPNPSDCDYNNILVLSYAAQANLHLTVLN 222
QY 239 QGAELADEWNADIHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKNLRDEP-- 296
DB 223 QAVKEAYLKNRRQFDYLEP-LPTAIDYVPVLTAKIEDYNYCVTVTKYKGLNLKTTTDS 281
QY 297 ----NMKWSIFNDYRYMTITVLDTISQPSLYDIKRYRDSIGGIEVKIGKNELTREIYTT 352
DB 282 NLDGNINWNTYRTKMTTAVLDLVALFPNDVGKY-----PIGVQSELTRREIYQV 333
QY 353 EINFDRLPOLRVQPNLATMEYNLTRASFLEQFIY-----TENTNFGNR----- 401
DB 334 -LNFESPYKYY--DFQYQEDSLTRRP-HLFTWLDLSLNFYEKAQTTPNPFTHYNNFHY 389
QY 402 -LVGISNRDAPYSNTITETLYGERGTSPTTKTIRPFESYKYSIVTDROSPPVSPQPHF 460
DB 390 TLDNISQSSVFGNHNVTDKL--KSLGLATNIYI-----F 422
QY 461 IINQIEL---YLNQSSNNT-----LKYSAGGSLSNYQ-NTTFFQPRKKDC 502
DB 423 LLNVISLDNKNYLDNYSKMPFFITNGTRLEKELTAGSGQITVDYVKNKIFGLPKRR 482
QY 503 NLVIDPGCSNPNFNYSKSHLSHSLFTYSYVIGIQLQILDGTGLVGTWTHSSVDYRNAISDKI 562
DB 483 ENQGNPTLPFTYDNYSHILSFILKSLSPATYKTQVY-----TFAWTHSSVDPKNTIYTHL 537
QY 563 ITMIPAIKGNLDTNSKVIETGPGHTGCGNLYVLOSQGRLEITCETPNSTQSYFIRLYATN 622
DB 538 TFOIIPAVKANSIGTAKSVQVGGPGHTGGDLI--DFKDHFKITQHSFNQFSYFIRIYASN 595
QY 623 GAGNTLPNTSLTPGVIGIPQRLNNTFSGTNNLYQVDFGVGFQFPSTVTLPLNRPNIP 682
DB 596 GSANTRAVNLISIPGVAEL--GHALNPTFGSDTNNLYKYKDFQYLEFNEVKFAPNQISL 654
QY 683 IFNRADV--SNSLIIDIKBIFIPITSSMQNREKQKLETIQTKINTFTNHTK 733
DB 683 IFNRADV--SNSLIIDIKBIFIPITSSMQNREKQKLETIQTKINTFTNHTK 733
```



```
DR GO: 0030435; P: sporulation; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR Pfam; PF03944; Endotoxin_C; 1.
DR Pfam; PF00555; Endotoxin_M; 1.
DR Pfam; PF03945; Endotoxin_N; 1.
SQ SEQUENCE 675 AA; 77759 MW; 4EB0E51AA0372FF1 CRC64;

Query Match      23.1%; Score 892.5; DB 2; Length 675;
Best Local Similarity 31.9%; Pred. No. 4e-49;
Matches 245; Conservative 126; Mismatches 267; Indels 129; Gaps 30;

QY 1 MNQNDNNYEIDSHSTPYFPNRRNSDGRYPYNNPNQPLQNTYKEMLMQCGNTQYG 60
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 1 MNPYQNKNEYEFNAPSNGF--SKSNYSRYPLANKPNQPLKNTYKDWLVNVCQDNOQYG 58
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||

QY 61 DNFEFPASADTTAAVSAGTIVSTGLAGIGLTSISGPIGIIAIIISFGTLITVFWPAG 120
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 59 NNAGNFPASSETIVGSAGIIVVGTMLG-----AFAAP--VLAAGIISFGTLPIFW-QG 109
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||

QY 121 EQDKTWTQIKMGEIFVDTPLTE---SIKQLKLTLEGFRQLQSYNTALDDWRKLR 177
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 110 SDPANVQDLNIG-----GRPIQEIKNIIIVLTSITVPIKNQDKYQEFDFKWEPA 163
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||

QY 178 QAPGLPPSSALQQAALTKIRFENVH---DFIREIPGFOLEYTKTLLPIYAQAANFHL 234
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 164 -----THANAKAVHDLFTLEPIIDKOLDMLKNNASYRIPT-----LPAYAIATWHL 211
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||

QY 235 NLLQGAELADWADIHPSQIEPNAGTSDDYK-LLENIPKYSNCANTYRTGLKNLR 293
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 212 NLLKHAATYNIW---LQNGINPSTFNSNYYQYGLKRGKIQEYTCIQTYNAGLTMR 268
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||

QY 294 DEPNKMSLFNDYRMTITVLDTSIQFSLYDIKRYRDSIGGIEVGKIKNELTRELYTTE 353
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 269 TWTNATWNYITRYLEMTITVLDLIAIFPNYDPEK-----PIGVKSELIREVY-TN 319
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||

QY 354 INFDRLPQLRVQPNLATMEYNLTRASFLEQFIFVTENTN-----FGNRL 402
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 320 VNSDIF-----RTILENGLTR-NPTLFTWINGRFFYTRNSRDILDPYDIFSFTGNQM 372
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||

QY 403 VGISNRDAPTYNTITETLYGERTG---SPTTKTIRPFESY-----KVSIVTDRQSPVPSP 455
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 373 -----APHTNDRDRIIAGVHGNIIISQDTSKVFPFVRNKPDKVEIVRHREYSDI 424
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||

QY 456 IQPHFIINOELYLNGSSNNTLYKSAGSL-SNYQNTTFQPRKKDCNLVIDPGSPNF 514
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 425 -----YEMIFFNSSEVFYSNNTENNYKRTDSTMPKQTKWN----- 464
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||

QY 515 NNYSHLSHSLFTYSYVIGLQIOLDTGVLGWTHSSVDRYNAISDKIITMIPAIGNNL 574
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 465 KEYGHTLSYIKTDNYIFSUVRRRV---AFSWHTSVDFQNTIDLDNITQIHAKALKV 521
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||

QY 575 DTNSKVIQPGHTGGNLVYQSGRLIEITCETPNSTQSYFIRLRYATNGAGNTLPNISLT 634
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 522 SSDSKIVKPGHTGGDLVILKDSMDPRVF-LKNVSRQYQVRIYATNA-----PKTTFV 575
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||

QY 635 IGVGIGIPQRLNNTSGTNYN--NIQYGDGFGYQFPSTV-----TLPLNRNIPFI 683
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 576 LTGIDITISVE-LPSTTSRQPNATDITYADFGYVTFPTVPNKTFEGEDFLMT----- 629
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||

QY 684 ENRADVSNILIDKIEFTPTSSMHONREKQKLEITQTKINTFFTN 730
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 630 YGTPNHSYNI-YIDKIEFIPITQSVLDYTEKQNIETKQKIVNDLFVN 675
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||

RESULT 5
Q9FDC0 BACTF PRELIMINARY; PRT; 1128 AA.
AC Q9FDC0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
```

Db	189	WNADQHPSPMLKSGT---	YYDELLVYIEKYINYCTKTYHKLGNHLKGSEKITWDAYNTY	244
Qy	307	RRYMTITVLTISOFSLYDIKRYRDSIGGI	EVKGIKNELTREIYTFEINFDRLPQLRVQP	366
Db	246	RREMTLVLDVAIVFPFVIDIRF-----	PRGVELELTREYVTSLDHLTRPP-----	291
Qy	367	NLATWEYNLTRASPKLFSFLBQFI	FYTYTNTWFGNRLVGISNRDAPYSNT---	423
Db	292	-----GLFTWLSDI	ELYTESVAEGDYLSGI--RESKYTCNQPF	335
Qy	424	ERTG-SPTTKTIRPESKYSIVTRQSP	PPVSPLOPHFIINQIELYLANGSSNNTLKYSAG	482
Db	336	NTNRLSKQLITLLPGE-FWTHLS	INRPPTTAGINKLYSLQIKIVFTTFKNDN--	392
Qy	483	GSLSNY---QNTTFQFP	PRKKDCNLVIDPGSGSPFNNTSHLISHFSLFTYS	539
Db	393	FVNUNQNEPQETT-----	NYPNDYGGG-NSQKFKNL	435
Qy	540	LD-----TGVLGWTHSV	VDRYNAISDKIITMTIPAKIGNNLD	595
Db	436	AEYFHSIFALGWTNSVNSQNLIS	ESVSTQIPLKAYEV-TNNSVIRGPGFTGGDLIELR	494
Qy	596	SQGRLEITCETPNSTQSVF	RLRYATGAGNTLNPISLITIPQVIGIPQRLNNTFS	655
Db	495	D-KCSIKCKA-SSEKKY	AIASLSFVAANNAIAVSDVGSAGVL-----	545
Qy	656	NN-----LOYGDFCY	FQFPSTVTLPLNRNIPFI	708
Db	546	NNFTIQDLNKYDFQYHTLL	VDIELPESEIHHLKREDDYEEGVLLIDKLEFKPIDENY	605
Qy	709	HQNRKQKLEITQTKTFF	TNHTKTL 735	
Db	606	---TNEMNLEKAKK	AVNVLFINATNAL 629	
RESULT 7				
Qy	Q8VNX2	BACTV		
Id	Q8VNX2	BACTV	PRELIMINARY; PRT; 650 AA.	
AC	Q8VNX2			
DT	01-MAR-2002	(Tremblrel. 20, Created)		
DT	01-MAR-2002	(Tremblrel. 20, Last sequence update)		
DT	01-MAR-2004	(Tremblrel. 26, Last annotation update)		
DE	Cry29Aa	protein.		
GN	Name=cry29Aa;			
OS	Bacillus thuringiensis	(subsp. medellin).		
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;			
OC	Bacillus cereus group.			
OX	NCBI_TaxID=79672;			
RN	[1]			
RA	NUCLEOTIDE SEQUENCE.			
RA	Delecluse A., Orduz S.;			
RL	Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AJ751977; CAC80985.1; -; Genomic_DNA.			
DR	GO; GO:0016787; F:hydrolase activity; IEA.			
DR	GO; GO:0005102; F:receptor binding; IEA.			
DR	GO; GO:0006952; P:defense response; IEA.			
DR	GO; GO:0009405; P:pathogenesis; IEA.			
DR	GO; GO:0030435; P:sporulation; IEA.			
DR	InterPro; IPR001178; Endotoxin.			
DR	InterPro; IPR005638; endotoxin_C.			
DR	InterPro; IPR005639; endotoxin_N.			
DR	Pfam; PF03944; Endotoxin_C; 1.			
DR	Pfam; PF00555; Endotoxin_M; 1.			
DR	Pfam; PF03945; Endotoxin_N; 1.			
SQ	SEQUENCE 650 AA; 74435 MW; DOCB846377CC517B CRC64;			
Query Match 21.9%; Score 847; DB 2; Length 650;				
Best Local Similarity 30.5%; Pred. No. 3.4e-46;				
Matches 230; Conservative 124; Mismatches 273; Indels 126; Gaps 22				
Qy	1	MNQNDNNEVEI	IDSHTSPVFPFRNSNDSRYPTYNPNQPLONTNYKEWLNC--	QGNVQ 58
Db	1	MNFGQNKNEY	ILNAPGS--ISNTPNNYSKYPIANTLNTQNTNTQNTNTKCAIDNNLK	58

Qy	59	YGDNFEFPASADTTAAVSAGTIVSGTTLIAGIGGLTISISGPIGIIIGAIISFQTLITVFWP	118
Db	59	SINPFEANLQNSLVGLFAITAAIASLLSAPITGGTSAAGTAIAAAI-----IPILWP	111
Qy	119	AGEQDKTWTQFIKNGELFVDTPLTSESIKOLKQLQLEGPQIQLQSNTALDWRKLRQL	178
Db	112	S--QBNLPLDKLLAISEATLYSFQDQVRVEDALTRESLKSQVKFENAFFTWIN-----	164
Qy	179	APGLPPSALQAAALTAKIRFENVHNDFIREIPGFOLEYTKTLPLPIYAQAANFHLNLQ	238
Db	165	----NPNST--NTTWRERFQEVNGRFVGSMAFPRAKNYBPIILLSTYAQAARLHLHLR	217
Qy	239	QCAELADEWNADIIHPSQIEPNAGTSDYYKLLKENIPKYSNYCANTYRTGLKNLRDEPNM	298
Db	218	DGITYAEKWNLSROGDDM-----PGDLLYKGFNKYCNEYIEHCIKWYNESLSLK--SVGA	271
Qy	299	KWSIFNDYRRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKGIKNELREIYTTTEINPDR	358
Db	272	NWLEYNQYRTFLTASVLDVLSFSDYDRLKYERL--SVEI-----LTRKLYDPINVHR	324
Qy	359	LPQLRVOPNLATMEYNLFRASFKLPSFLEQFIYFTENTFNGRLVGISNRDAPTYSNIT	418
Db	325	GISLEADESKYTLEPTLFTQLYVTLTFYSNIFVNYMGHTN-----TYRYLSP	370
Qy	419	ETLYGERTGSTTKTIRPPESYKYSIVTDROSPVPSPQPHIINOIELYLN-----	470
Db	371	DKIPAERSFGQSSYID-----KVPVIDNKS-----IYKIRAYDNHNGLENVM	415
Qy	471	-----GSSNNTLKYSAGGSLSNQNTTFFQPRKKQCNLVDPGCSFNFNYSHLHS--	523
Db	416	YFGWDEKDOIQLIGSSSTEY-----IKNCT-----HRLADVISHD	454
Qy	524	--FSLFTYSYVIGLQLILDTVGLGTHSSVDRYNAISDKIITMIPAIGNNLDTSKVI	581
Db	455	LDEKNKCYSF-----AWTSTTISLENEIKNDIITQIPAVKAYQLGVQSQVI	500
Qy	582	EGPHTGGLNLYVLOSQLEITCE--TPNSTOSYFIRLRYATNGANTLPNISLTTIPGWIG	640
Db	501	KGPQHTGDDLNLKNSDNYLRISCOHLSNVTKKYFVRIRYATNGSLNTRPIINITP--G	557
Qy	641	IPPOR--LNNTFSGTNYNLYQDPGYQFQFSTVPLNLRNPIPFENRADV--SNSILIID	697
Db	558	MTPQGMVLNDFTSCTGYSNLYQNGFYKEFLKEVTLNPNQISLSLTNRSQNSIILLD	617
Qy	698	KIEFIPITSSMHQNRKQKLETIOTKINTFTFN	730
Db	618	RIEFLPITPSIRKSGKEONLEKNOKTVNKLFPN	650

```

RESULT 8
Q75VA2_BACTE
ID Q75VA2_BACTE PRELIMINARY; PRT; 683 AA.
AC Q75VA2_
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative mosquitoicidal toxin.
GN Name-cry30Aa like;
OS Bacillus thuringiensis (subsp. entomocidus).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=1436;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=INA288;
RA Ikeda T., Yamaya K., Ito T., Sahara K., Bando H., Asano S.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB125059; BAD00052.1; -; Genomic DNA.
DR GO: 0016787; F:hydrolase activity; IEA.
DR GO: 0005102; F:receptor binding; IEA.
DR GO: 0006952; P:defense response; IEA.
DR GO: 0009405; P:pathogenesis; IEA.

```

```

DR GO: GO:0030435; P:sporulation; IEA.
DR InterPro: IPR001178; Endotoxin.
DR InterPro: IPR005638; endotoxin_C.
DR InterPro: IPR005639; endotoxin_N.
DR Pfam: PF03944; Endotoxin_C; 1.
DR Pfam: PF00555; Endotoxin_M; 1.
DR Pfam: PF03945; Endotoxin_N; 1.
DR SQ SEQUENCE 683 AA; 77438 MW; 488B477B876FDCA CRC64;

Query Match 20.8%; Score 803.5; DB 2; Length 683;
Best Local Similarity 31.6%; Pred. No. 2.4e-43;
Matches 244; Conservative 113; Mismatches 284; Indels 131;

Qy 1 MNQNNDNNEYIIDSHTSPYFPNRSNDSRYPTNPNOPLOQNTNYKELWLNCCQGN
Db 1 MNPYQKTEYEILDA-LPNYSNMVNAYSRYPLANNPQVPLQNTSYKDWLANWCQ-
Qy 61 DNPETFASADT-----IAVSAAGTVSGTLAGIGGTSISGPIGIIIGAIISFGT
Db 53 ----TINPLCTPINIDISLAASIAVVGSTLA-----LIPGGEAIGFVLGTFTST
Qy 116 FWPAGEQDKVTWTFIKMG-EIFVDTPLTESI KQ LK LQ TLEGFRQI---LQSYNTFA
Db 103 LMPNGE--TKIWDTPAERGLQFPELPGQDAIEILTVGVKSGYNALKNRMENFEQQA
Qy 172 RKLKELQAPGLPPSSALQQAALTLKIR-FENVHNDFREIPGQLE-TYKTELLPLI
Db 161 K-----GNRTSNAEQV-----IRDFQSDVKDILKNDYMINPENKPAFIML
Qy 230 ANPHNLQQOAEALADEWNADIHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYY
Db 208 ANIDLILYQGA VYGDWDEKDINGS-ISPFWG-SKDYESLTKIEEYTYNYCAEY
Qy 290 KNLDEPNKWSIFNDYRRYWTITVLDTISQFSLYDIKRYRDSIGGIEVKGIGKLEL
Db 266 NILKNKPNISDWTYKNYRREATLGALDLVAFPNYDMHLY-----PAATKTESEL
Qy 350 YYTEINFDRLPQLRQVPNLATMEYNLTRASFCLKFSLEQFI FVTNTWFG-----N-N-N-N
Db 318 YMPSPG---LQSNVFSQLEGENALTHPP-SUFTWLNELNLYTVRENFPALQVSI
Qy 405 ISNRDAPTYSNITITET-LYGERTGSPTTKTRIRPFSYKVSIVTDRQSPVSPVSIOPH
Db 374 LQARSRYTONPTILDNPAQGVNRGTSQIGLNNLFVYKLSMSQYHHHPNDCSSIAG-
Qy 464 QIELY---LNGSSNNTLYKSAGGSLSYQNVTTFQFPRKDCNVLVDPCSSPNFNNN
Db 431 DMTFYKSDYNGNAGATQTYOAGRNTNNVIN-TEWNGPQK-----ASSSSNNN
Qy 519 ----HILSHFSLFTYSYVIGLQLDLDTGV-LGHTHSSVDRYNAISDKITIMIPATK
Db 479 QTKHLSDIKMIYFR--TGGMYQYDFGYSFAWHTSVDPDNLIVENRITQIPAVKY
Qy 575 DTNSKVIISGPGHTGGLNLYLSQSGRLBITCETPNSTQSFIRLRVATNGAGNT----
Db 537 STDSKVGKVPFVGDDLKLLKQATIRI--KTDHANTRYKIRVRASN--ANTPLV
Qy 629 PN-ISLTIPGVIGIPORLANNTPFSGTNNYNNLQYDGFYGFPTSTVTLPLNRNIPPI
Db 593 QNTLIVTFP-----QTITHSTISELQYKDFQYVTFPG-----EPI
Qy 688 DVNSGI-----LIIDKIEFTIPTSSMHQNRKQKLETTQTKINTFTTN 7300
Db 632 SIDVAIRGVONDRINDIWRLEFPIPTQSVDLYTEQNIEKSOXKAVNDLFIN 683

```

RESULT	9
Q8VNX1	BACTV
ID	Q8VNX1 BACTV PRELIMINARY; PRT; 688 AA.
AC	Q8VNX1;
DT	01-MAR-2002 (TREMBlrel. 20, Created)
DT	01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT	01-MAR-2004 (TREMBlrel. 26, Last annotation update)


```
DE Cry30Aa protein.
GN Names=Cry30Aa;
OC Bacillus thuringiensis (subsp. medellin).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=79672;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Delecluse A., Orduz S.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL, A2251978; CAC80986.1; -; Genomic DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0005102; F:receptor binding; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR GO; GO:0030435; P:sporulation; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR Pfam; PF03944; Endotoxin_C; 1.
DR Pfam; PF00555; Endotoxin_M; 1.
DR Pfam; PF03945; Endotoxin_N; 1.
SQ SEQUENCE .688 AA; 77844 MW; 75057085BD7B3A9 CRC64;

Query Match 20.3%; Score 784; DB 2; Length 688;
Best Local Similarity 32.1%; Pred. No. 4.4e-42;
Matches 237; Conservative 113; Mismatches 294; Indels 94; Gaps 30;

QY 1 MNQNDNNEYIIDSHTSPYFPRNSDRSYPTNNPQNPLQNTYKRWLNMCQNTQYG 60
DB 1 MNSYKNKEVEILD--TSQKNSMNSRYSKYPLTNPKVPLQNTYKDWLNMCQTITPLC 58
QY 61 DNPETASADTIAAVSAGTIVSGTLLAGIGLISGPIGIGIAIISPGTLITVFPAG 120
DB 59 TPIDTDSKLVAIKVIGAI-----FKSMGPGGAAGVLKSFSTIIPILWP-- 105
QY 121 EODKT-VWTFQIKWG-EIFVDTPLTESIK-----QLKLTLEGFRLOLSYNTALDWR 172
DB 106 -NDKTPWKEFTQGLQFLPELGRDAIEIIGNDVQAEYNSLE---IMMRDFENKFW 161
QY 173 KLRQLQAPGLPPSSALQQAALTILKIRFENVHDFIREIPQFL-ETVKTLLEPIYA 231
DB 162 -----SNRTFANAIAVTAFTSTVNTQIIRKFLAPENRPAFLNLYAQTAN 209
QY 232 FHLNLLQQAELADEWNADHPQSIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKN 291
DB 210 IDILLYQGSVYGKQWADINNRSTSP--PSSKDYQSLGKIKDYKTYNCAETVRSNLT 267
QY 292 LRDEPNMKHSIFNDYRYMTITVLTISQPSLDIKRYRDSIGIEVKGINKELTREIYT 351
DB 268 LKMKPHIQWDIYRNYRREALGALDLVALPFPNYDICIYPTQ-----TRTELTRK 319
QY 352 TEINPRLPQLRQPNLATWEYNLTRASPKLPSFLEQIFITYENTNFGN--RLVGISNRD 409
DB 320 PSFYLAQLQORDIE-----TVENQLTHPP-SLFTWLNLNLYTTRERPNVLQVASL 374
QY 410 APT-----YGNITETILYGE-RTGSPPTTKTIRPFESYKSVISVTRDQSPVPSP 464
DB 375 ATSRYTQNTTISNPVQGPVREGTPTKISLANYIYKLFMSQYRHPNDCLPISG--- 431
QY 465 IEL-----YLGSSNNTLYKYSAGSLSNYQNTTFFQPRKKDCNLVIDPGCSPN 520
DB 432 MSFYRSDYDYGAGPAPVHYSGASPTNVIK-TYMNGPQ-----NALISNDIS-- 484
QY 521 LSHFSLFTSYVIGLQQLIDTGV-LGWTHSSVDYRNAISDKLITMTPAKGNLDTNSK 579
DB 485 LSDIKM-NYSRTGCV-YPLYDFGYSFAWTHTSVDDPNLIVPNRITQIPAVKASLTSP 542
QY 580 VTGPGHTGNLVL-----OSQGRLEITCTPNST---QSYFIRLRVATGAGNLTLP 632
DB 543 VIVPGHTGDDLVALNSGTQSTQMTQCKTGFTGFSRQYGLRMRYAANSATV--SL 600
QY 633 LTPGVIGIPQRLNNTFSGTNY---NNLQYDGFYGFQFPS--TVTLPLNRNPIPF 687
```

```
DB 601 YTLGQRTGTSFVTEFSRNIIPTDLKYEFKYQDYLQIITWLTPLNTITISMQQA 660
QY 688 -DVNSNLIIDKIEPIPI 704
DB 661 TGLLNQLIIDRIEFPYM 678

RESULT 10
CR8BA_BACUK
ID CR8BA_BACUK STANDARD; PRT; 1169 AA.
AC Q45705;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Pesticidal crystal protein cry8Ba (insecticidal delta-endotoxin
DE CryVIII(a)) (Crystalline entomocidal protoxin) (134 kDa crystal
DE protein).
GN Name=cry8Ba; Synonyms=50C(b), cryVIII(a);
OS Bacillus thuringiensis subsp. kumamotoensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
NCBI_TaxID=132267;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=NRRL B-18746 / PS50C;
RA Michaels T.E., Fonceirada L., Narva K.E.;
RT "Process for controlling scarab pests with Bacillus thuringiensis
RT isolates.";
RL Patent number WO9315206, 05-AUG-1993.
CC -!- FUNCTION: Promotes colloidsmotic lysis by binding to the midgut.
CC epithelial cells of insects. Active on various scarabaeid beetles.
CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; U04365; AAA21118.1; -; Genomic DNA.
DR HSPP; P07130; 1DLCL.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR Pfam; PF03944; Endotoxin_C; 1.
DR Pfam; PF00555; Endotoxin_M; 1.
DR Pfam; PF03945; Endotoxin_N; 1.
KW Sporulation; Toxin.
SQ SEQUENCE 1169 AA; 133544 MW; 22BFCF5BD699909 CRC64;

Query Match 20.0%; Score 775.5; DB 1; Length 1169;
Best Local Similarity 28.5%; Pred. No. 3.3e-41;
Matches 217; Conservative 136; Mismatches 289; Indels 119; Gaps 27;

QY 5 NDNNEYIIDSHTSPYFPRNSDRS-RYPYTNPNQPLQNTYKRWLNMCQNTQYGDN 62
DB 4 NNQNEYIIDSHTSPYFPRNSDRS-RYPYTNPNQPLQNTYKRWLNMCQNTQYGDN 59
QY 63 FEFPASADTIAAVSAGTIVSGTLLAGIGLISGPIGIGIAIISPGTLITVFPAGSQ 122
DB 60 PEVFLSQD--AVKAAIDIVGKLTGLG-----VPFVGPIVSLYTLQILDILMPS--K 107
QY 123 DKTWTQTFIKMGEIFVDTPLTESIKQLKLTLEGFRQLQSYNTALDWRKLRQLQAPGL 182
DB 108 QKSQWEIFMQVEELINQKIAEYARNKALSLEGLGNLYQLYLTALKEWKE----- 158
QY 183 PPSALQQAALTILKIRFENVHDFIREIPGQLETYKTLTLLPIYAAANFHLNLLQQGAE 242
```

Db 159 NPGS--RALRDVNRPEILDSLFTQYMPSTRVNFVPELTYTMAANLHLLLRDASI 216
Qy 243 LADEWADIHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKNLRDEPNMKWSI 302
Db 217 FGEEWGL-----STSTINNYNRQMKLTAEYSDHCVKWYETGLAKLGSSAKQWID 267
Qy 303 FNDYRYWITVLDTSQSFLYDIKRYRDSIGGIEVKGINKELTREIYTEINFDRLPOL 362
Db 268 YNQFREMTLVLDVVALFSNYDTRY-----PLATTAQLTREYVYTDPLGAVDVPI 319
Qy 363 RVQPNLATMEYNLTRASFK---LFSLEQPIFYTENTNF-----GNRLVIGISNRD 409
Db 320 GSWYDRAPSEFEKAAIRPPHVDYITGLTVYTKRSFTSDRYMYWAGHQ---ISYKH 376
Qy 410 APTYSNTITETALYGERGTGPTTKTIRPFESYKV-----SIVTDROSPPVS-----PI 456
Db 377 IGT-SSTFQ-MYGTQNLQSTSNF-DFTNYDIYKTLNGAVLIDVYFGYTFYFPGMPE 433
Qy 457 QPHFIINQIELYNGSSNNTLKYSAGSLSNYQNTFFQPPRKQCNLVIDPOCS--PNF 514
Db 434 TBFFMVNQL-----NNTRK-----TLTYKPASKDIIDTRDSELELPETSGQPNY 479
Qy 515 NNYSHILSH-----FSLFTSYVYGLQLOLDTGVLGTWTHSSVDRYNAISKIITMIPA 570
Db 480 ESYSHRLGHITFYSSSTSYV-----PVFSWTHRSADLTNTVKSGETIQIPGGK 529
Qy 571 GNNLDTNSKVGEPGHTGNNLVYLSQO-GRLEITCETPNSTQSVFIRLRYATNGAGNTLP 629
Db 530 STIGRNTYIKRGYTGGLVALTRIGSCFQMIPPES-QRFRIRIRYASNETS-----584
Qy 630 NISLTIPGVIGIPPORLNNTFSCTNNYNNLQYDGFYQFPSTVTLPLNRRNIPPIFNRA 689
Db 585 --VISLYGLNQSTLKFNTQYSKNENDLTYNDFKIEYPRVSNASSNIQRLSIGIQ 642
Qy 690 SNSIILIDKIEPIPTSSMHQNRKQKLETIQKINTFFTN 730
Db 643 NTNLFILDRIFIPVDTEYAEATD---LEAKKAVNALEFTN 680

RESULT 11
CR4BA_BACTI STANDARD; PRT; 1136 AA.
AC P05519; P11782; P16479;
DT 01-NOV-1988 (Rel. 09, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Pesticial crystal protein cry4Ba (insecticidal delta-endotoxin
DE CryIVB(a)) (Crystalline entomocidal protoxin) (128 kDa crystal
DE protein).
GN Name=cry4Ba; Synonyms=bt8, cryD2, cryIVB(a), isrH3;
OS Bacillus thuringiensis subsp. israelensis.
OC Plasmid 72 Kb.
OG Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=1430;
RN [1]_TaxID=1430;
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=88157738; PubMed=281510;
RA Tungpradabkul S., Sattasatian C., Panyim S.;
RT "The complete nucleotide sequence of a 130 kDa mosquito-larvicidal
RT delta-endotoxin gene of Bacillus thuringiensis var. israelensis";
RL Nucleic Acids Res. 16:1637-1638(1988).
RN [2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=4Q2-72;
RX MEDLINE=88185334; PubMed=2833395;
RA Chungjatupornchai W., Hoesfte H., Seurinck J., Angsuthanasombat C.,
RA Vaek M.;
RT "Common features of Bacillus thuringiensis toxins specific for Diptera
RT and Lepidoptera";
RL Eur. J. Biochem. 173:9-16(1988).
RN [3]

RP NUCLEOTIDE SEQUENCE.
RA Sen K., Honda G., Koyama N., Nishida M., Neki A., Sakai H., Himeno M.,
RA Komano T.;
RT "Cloning and nucleotide sequences of the two 130 kDa insecticidal
RT protein genes of Bacillus thuringiensis var. israelensis";
RL Agric. Biol. Chem. 52:873-878(1988).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=88329719; PubMed=2901387; DOI=10.1016/0378-1119(88)90229-6;
RA Yamamoto T., Wackinson I.A., Kim L., Sage M.V., Stratton R.,
RA Akande N., Li Y., Ma D.-P., Roe B.A.;
RT "Nucleotide sequence of the gene coding for a 130-kDa mosquitoicidal
RT protein of Bacillus thuringiensis israelensis";
RL Gene 66:107-120(1988).
RN [5]
RP NUCLEOTIDE SEQUENCE OF 1-134.
RX MEDLINE=88038331; PubMed=2890080; DOI=10.1007/BF00328128;
RA Angsuthanasombat C., Chungjatupornchai W., Kertbundit S.,
RA Luxananil P., Sattasatian C., Wilairat P., Panyim S.;
RT "Cloning and expression of 130-kd mosquito-larvicidal delta-endotoxin
RT gene of Bacillus thuringiensis var. israelensis in Escherichia coli";
RL Mol. Gen. Genet. 208:384-389(1987).
CC -I- FUNCTION: Promotes colloidsmotic lysis by binding to the midgut
CC epithelial cells of mosquitos.
CC -I- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -I- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -I- SIMILARITY: Belongs to the delta endotoxin family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; X07082; CAA30114.1; -; Genomic DNA.
DR EMBL; X05692; CAA29174.1; -; Genomic DNA.
DR EMBL; D00247; BAA00178.1; -; Genomic DNA.
DR EMBL; M20242; AAA22337.1; -; Genomic DNA.
DR EMBL; X07423; CAA30312.1; -; Genomic DNA.
DR PIR; S00398; USBS81.
DR PDB; 1W99; X-ray; A=84-641.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR Pfam; PF03944; Endotoxin_C; 1.
DR Pfam; PF00555; Endotoxin_M; 1.
DR Pfam; PF03945; Endotoxin_N; 1.
KW 3D-structure; Plasmid; Sporulation; Toxin.
FT CONFLICT 51 51 V -> D (in Ref. 4).
FT CONFLICT 65 65 T -> S (in Ref. 4).
FT CONFLICT 193 201 LINAQWSL -> PHKTRMYV (in Ref. 4).
FT CONFLICT 203 204 RS -> C (in Ref. 3).
FT CONFLICT 205 206 AG -> C (in Ref. 4).
FT CONFLICT 205 205 A -> R (in Ref. 1).
FT CONFLICT 272 272 Y -> L (in Ref. 3).
FT CONFLICT 325 325 D -> Y (in Ref. 3).
FT CONFLICT 364 401 FGSNLTHQIQNSNVKTSITDTSPPSNRVTKMDFYKI ->
FT LVQYLIKFNILIVILKLSQILAPPLIELQKWISTKF
FT (in Ref. 4).
FT CONFLICT 467 467 K -> N (in Ref. 4).
FT CONFLICT 496 496 K -> R (in Ref. 4).
FT CONFLICT 519 519 G -> GG (in Ref. 4).
FT CONFLICT 551 572 LNVSYLQGVSRGTTISTESTF -> IECDHMYKEFLEEQ
FT RLQVNYV (in Ref. 4).
FT CONFLICT 594 594 F -> N (in Ref. 4).
FT CONFLICT 687 690 ELYP -> GIIS (in Ref. 4).
FT CONFLICT 721 721 A -> R (in Ref. 4).
FT CONFLICT 823 832 SNRCETSAVP -> LIIVSVCECA (in Ref. 4).
FT CONFLICT 836 837 GN -> WD (in Ref. 4).
FT CONFLICT 902 902 E -> R (in Ref. 4).

FT CONFLICT 1015 1015 G -> V (in Ref. 4).
SQ SEQUENCE 1136 AA; 127764 MW; 8AC4E8C26FE3E9B5 CRC64;

Query Match 19.6%; Score 758.5; DB-1; Length 1136;
Best Local Similarity 31.1%; Pred. No. 4e-40;
Matches 232; Conservative 123; Mismatches 266; Indels 125; Gaps 33;

QY 28 DSRYPYTNPNQPLONTNYKEWLNMCQNTQYGDNFETFAADTIAAASAGTIVSGTLA 87
DB 2 NSGYPLANDLQGSMMKNTNYKDWLACENNQQYGVN----PAINSSSVSTALKVAGAI-- 55
QY 88 GIGGLTSISGPIGIIIGAIISFGTLITVFWPAG-EQDKTVMTQFIKMGIEIFVDTPLETESI 146
DB 56 ----LKFVNPAGVTLTVL---SAVLPLMPTNTPTPERVWDFMTGNTGLIDQTVTAYV 108
QY 147 KQLKLTLEGFRQILQSYNTALDDWRKLRLOAPGLPSSALQQAALTILKIRFNVHNDP 206
DB 109 RTDANAKMTVVVDYLDQYTKFTWKR-----EPNNQSYRTAVIT---QF-NLTSAK 156
QY 207 IREIPGF--OLETVKTLPIYAQAANFHLNLLQGAELADEWNADIHPSQIEPNAGTSD 264
DB 157 LRETAVFSNLGVYELLPLPIYAQVANFNLLIRDLGLINAQEWS-----LARSAG--D 207
QY 265 DYYKLLKENIPKYSNYCANTYRTGLKNLRDPNNKWSIFNDYRYMTITVLDTISQFSLY 324
DB 208 QLYNTWQYTKETIAHSITWYKGLDVLNRKSNQGWITFNDYKREMTIQVLDILALFASY 267
QY 325 DIKY-RDSIGGIEVKIGKNELTREIYTHEINFDRLPQLRVQPNLATMEYNLATRASFKLF 383
DB 268 DPRYPADKIDNTKLS--KTEFTREIYTA-----LVESPSKSAIAALEAALTR-DVHLF 318
QY 384 SFLEQFIYFTENTFNGNRLVGSINRDAPTYSNIT---ETLYGER-TGSPPTTKTIRPFES 439
DB 319 TWLKRVDFT-NTIYQDLRFLSANKIGFSYTNSSAMQESGIYSGSGFGLNTHQIQLNSN 377
QY 440 -YKVSIVTDROSPPVSPQIHPFIINQIELYNGSSNTLK--YSAGGSLSNYQN----- 490
DB 378 VYKTSI-TDTSSP-----SNRVTKMDFYKIDGTLASYNSTIPT 415
QY 491 -----TTFOPPRKOCNLVIDPGCSFNFNYSIHLSHSLFTYSYVIGLQQLDITG-- 543
DB 416 PEGRLTTFPGFSTNEN-----TPN-QPTVNDYTHLSY-----IKTDVIDNSN 458
QY 544 --VLGTHSSVDRYNAISDKIITMIPAKGNLDNTSKVIEGPGHTGGNLVYLOS-----Q 597
DB 459 RVSFATWTKIVDPNNQIYTDAITQVPAVKSFNFLNATAKVIKPGCHTGGDLVALTSNGTLS 518
QY 598 GRLEITCEIP--NSTQSYFIRLRYATNGAGTILPNISLITPGVIGIPPORLNTTFSGTN 654
DB 519 GRMEIQCKTSIFNDPTRSYGLRIRYAANS--IVLNVSIVLQGVSRGCTTISTESTFSRPN 576
QY 655 Y---NNLQYGDGYFO-FPSTVTILPLNRN--IPFIFNRADV-SNSILIIDKIEPIPTSS 707
DB 577 NIIPDLKYEFRYKDPFDVAIVPMRLSSNQLITIAIQLPNTSNNQVIIDRIEIPITQS 636
QY 708 MHQNRKQKLETTQTKINTFTNHTK 733
DB 637 VLDETENQLESEREVNALFTNDAK 662

RESULT 12

Q7AL72 BACTI
ID Q7AL72 BACTI PRELIMINARY; PRT; 1136 AA.
AC Q7AL72;
DT 10-MAY-2005 (T-EMBLrel. 30, Created)
DT 10-MAY-2005 (T-EMBLrel. 30, last sequence update)
DT 10-MAY-2005 (T-EMBLrel. 30, last annotation update)
DE Pesticidal crystal protein cry4BA.
GN Name=cry4BA; Synonyms=Bt8, cryD2, cryIVB(A), isrH3, pBt038;
OS Bacillus thuringiensis (subsp. israelensis);
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=1430;

RN NUCLEOTIDE SEQUENCE.

RX MEDLINE=22323415; PubMed=12324359;
RA DOI=10.1128/AEM.68.10.5082-5095.2002;
RA Berry C., O'Neil S., Ben-Dov E., Jones A.F., Murphy L., Quail M.A.,
Harris D., Zaritsky A., Parkhill J.,
RT "Complete sequence and organisation of pToxis, the toxin-coding
RT plasmid of Bacillus thuringiensis subsp. israelensis";
RL Appl. Environ. Microbiol. 68:5082-5095 (2002).
DR EMBL; AL731825; CAD30095.1; -; Genomic DNA.
SQ SEQUENCE 1136 AA; 127763 MW; 8AC4E8C26FE3E9B5 CRC64;

Query Match 19.6%; Score 758.5; DB 2; Length 1136;

Best Local Similarity 31.1%; Pred. No. 4e-40;
Matches 232; Conservative 123; Mismatches 266; Indels 125; Gaps 33;

QY 28 DSRYPYTNPNQPLONTNYKEWLNMCQNTQYGDNFETFAADTIAAASAGTIVSGTLA 87
DB 2 NSGYPLANDLQGSMMKNTNYKDWLACENNQQYGVN----PAINSSSVSTALKVAGAI-- 55
QY 88 GIGGLTSISGPIGIIIGAIISFGTLITVFWPAG-EQDKTVMTQFIKMGIEIFVDTPLETESI 146
DB 56 ----LKFVNPAGVTLTVL---SAVLPLMPTNTPTPERVWDFMTGNTGLIDQTVTAYV 108
QY 147 KQLKLTLEGFRQILQSYNTALDDWRKLRLOAPGLPSSALQQAALTILKIRFNVHNDP 206
DB 109 RTDANAKMTVVVDYLDQYTKFTWKR-----EPNNQSYRTAVIT---QF-NLTSAK 156
QY 207 IREIPGF--OLETVKTLPIYAQAANFHLNLLQGAELADEWNADIHPSQIEPNAGTSD 264
DB 157 LRETAVFSNLGVYELLPLPIYAQVANFNLLIRDLGLINAQEWS-----LARSAG--D 207
QY 265 DYYKLLKENIPKYSNYCANTYRTGLKNLRDPNNKWSIFNDYRYMTITVLDTISQFSLY 324
DB 208 QLYNTWQYTKETIAHSITWYKGLDVLNRKSNQGWITFNDYKREMTIQVLDILALFASY 267
QY 325 DIKY-RDSIGGIEVKIGKNELTREIYTHEINFDRLPQLRVQPNLATMEYNLATRASFKLF 383
DB 268 DPRYPADKIDNTKLS--KTEFTREIYTA-----LVESPSKSAIAALEAALTR-DVHLF 318
QY 384 SFLEQFIYFTENTFNGNRLVGSINRDAPTYSNIT---ETLYGER-TGSPPTTKTIRPFES 439
DB 319 TWLKRVDFT-NTIYQDLRFLSANKIGFSYTNSSAMQESGIYSGSGFGLNTHQIQLNSN 377
QY 440 -YKVSIVTDROSPPVSPQIHPFIINQIELYNGSSNTLK--YSAGGSLSNYQN----- 490
DB 378 VYKTSI-TDTSSP-----SNRVTKMDFYKIDGTLASYNSTIPT 415
QY 491 -----TTFOPPRKOCNLVIDPGCSFNFNYSIHLSHSLFTYSYVIGLQQLDITG-- 543
DB 416 PEGRLTTFPGFSTNEN-----TPN-QPTVNDYTHLSY-----IKTDVIDNSN 458
QY 544 --VLGTHSSVDRYNAISDKIITMIPAKGNLDNTSKVIEGPGHTGGNLVYLOS-----Q 597
DB 459 RVSFATWTKIVDPNNQIYTDAITQVPAVKSFNFLNATAKVIKPGCHTGGDLVALTSNGTLS 518
QY 598 GRLEITCEIP--NSTQSYFIRLRYATNGAGTILPNISLITPGVIGIPPORLNTTFSGTN 654
DB 519 GRMEIQCKTSIFNDPTRSYGLRIRYAANS--IVLNVSIVLQGVSRGCTTISTESTFSRPN 576
QY 655 Y---NNLQYGDGYFO-FPSTVTILPLNRN--IPFIFNRADV-SNSILIIDKIEPIPTSS 707
DB 577 NIIPDLKYEFRYKDPFDVAIVPMRLSSNQLITIAIQLPNTSNNQVIIDRIEIPITQS 636
QY 708 MHQNRKQKLETTQTKINTFTNHTK 733
DB 637 VLDETENQLESEREVNALFTNDAK 662

RESULT 13

Q5W7N9 BACTU
ID Q5W7N9 BACTU PRELIMINARY; PRT; 688 AA.
AC Q5W7N9;

DT 01-FEB-2005 (T-EMBLrel. 29, Created)
DT 01-FEB-2005 (T-EMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (T-EMBLrel. 29, Last annotation update)
DE Cry30-like.
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=1428;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=96-OK-85-24;
RA Ohgushi A., Wasano N., Saitoh H., Ohba M.;
RT "Identification and characterization of novel cry genes from an
RL mosquito-specific Bacillus thuringiensis serovar sotto strain."
RL Submitted (Oct-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB193814; BAD67157.1; -; Genomic DNA.
DR GO; GO:016787; F:hydrolyase activity; IEA.
DR GO; GO:005102; F:receptor binding; IEA.
DR GO; GO:0003952; P:defense response; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR GO; GO:0030435; P:sporulation; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR Pfam; PF03944; Endotoxin_C; 1.
DR Pfam; PF00555; Endotoxin_M; 1.
DR Pfam; PF03945; Endotoxin_N; 1.
SQ SEQUENCE 688 AA; 77438 MW; D5F1C265AEC6BD1 CRC64;

Query Match 19.0%; Score 734; DB 2; Length 688;
Best Local Similarity 31.2%; Pred. No. 7.7e-39;
Matches 238; Conservative 111; Mismatches 271; Indels 142; Gaps 34;

QY 1 MNQNDNNEVEIIDSHTSPVFPNRNDS-----RYPVTNNPNQPLQNTYKELNWCQGN 56
DB 1 MDLYGNKNDSEILNASSN-----NSNMSTYKYLPLAHSRQDSMMENMYKEWINQC--- 51

QY 57 TOYGNFETPAS-ADT-IAAVSAGTIVSGTLGIGGLTSISGPFIGIGAILIISFGTLIT 114
DB 52 ----EFINTFCTPIDIDNSVAATIGAVGAILA-----LIPGGEAIGFVLSTFTSLIP 101

QY 115 VFWPAGEQKVTWQIKMG-ELFVDTPLTESIK-----QLKQTLLEGPRQILQSYNTA 167
DB 102 YLWPS-----DTKKMGDTQKGLQFLPELGNDAIEIIGNDVQSEYNSLKT-----MQNFEDS 156

QY 168 LDDWRKLKELQAPGLPPSALQQAALTAKIRFENVHDFIREIPGQLE-TYKTLPLPIY 226
DB 157 FTDWKKYRN-----RATAVAVTNDFSVRDQIIRLKDRLINPENKPAFLILY 204

QY 227 AQAAFHLLNLOQGAELADEWNADIHPSQTEPNAGTSDDYYKLLKENIPKYSNYCANTYR 286
DB 205 AQTANFDLILYQRGALYADEWENDINRS-ISPLLG-SKDYIISLAAKIKETNYCAETVR 262

QY 287 TGLKNLRDENMKWSIFNDYRYMTITVLDTSIQFSLYDIKRYRDSIGGIEVGKIKELT 346
DB 263 NSLNLKNTNISWGTNYKRYRREVTGLDLAALFPNYDICIPIQ-----TKTELT 314

QY 347 REIYTEINFDRLPQLRVQPNLATMEYNLTRASFLEQIFFTENTNFGNRLVGIS 406
DB 315 RKVYMPSFG---LQQSNYFQSLGLEGNALTHPP-SLFTWLNELNLYTIRENFNPALLVSS 370

QY 407 NRDAPTYG-----NTITELYGERGSPPTTKTIRFPESYKSVIVTDQSPVPSPQHP 460
DB 371 LSGLOAISRYQNPWRINSPAQVRNGTPTQIGLNNLFVYKLSL-----SQYHHPNICY 425

QY 461 I--INQIELY---LNGSSNNLTYSAGSLSNYQNTTFFQFPRKDCNLDVPCGSPNFN 515
DB 426 IAGISDMTFKSDYNGNAPTQYQAGRSNNFIN-TERKNGPQE-----ASSSN 473

QY 516 NYS-----HILSHFSLF-----TY-SYVIGLQQLDITGLVLTWTHSSVDRYNAISDKII 563
DB 474 NISIKQTNHILSDIKMIYRGTGTYPSPYDFGYS-----FAWTHTSVNPDNLIVPNRI 525

QY 564 TMIPAIGNNLDNTNSKVIIEGPGHTGGMNLVYL-----QSQGRLEITCETPNST---OSYFIR 616
DB 526 TQIPAVKADYLTSPAKVIAGFTGGDLVALLNAATQAGRMQIQCKTSGFTGASRRYGR 585

QY 617 LRYATNGA-----GNTLPNISITIPGVIGIPQRLNNTFSGTNNYNNLQYGDGFGY 665
DB 586 IRYAANALTVSLSYTVQGGNTMTTTFITERFTL-----RPNTTIP-----TDLKYEEFPKY 636

QY 666 FQFPPS--TVTLPLNRNIPFIFNRADV--SNSILIIDKIEFIPI 704
DB 637 KEYNQIITMAPQNTIVTIALQQLNAPFNQDLIIDRIEFYPM 678

RESULT 14
C19BA BACUH
ID C19BA_BACUH STANDARD; PRT; 682 AA.
AC O86170;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Pesticidal crystal protein cry19Ba (Insecticidal delta-endotoxin
DE CryXIXB(a)) (Crystalline entomocidal protoxin) (78 kDa crystal
DE protein).
GN Name=cry19Ba; Synonyms=cryXIXB(a);
OS Bacillus thuringiensis subsp. higo.
OG Plasmid.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=132266;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=98369730; PubMed=9704107;
RA Wang S.H., Saitoh H., Mizuki E., Higuchi K., Ohba M.;
RT "A novel class of mosquitoicidal delta-endotoxin, Cry19B, encoded by a
RT Bacillus thuringiensis serovar higo gene."
RL Syst. Appl. Microbiol. 21:179-184(1998).
CC -!- FUNCTION: Promotes colloid osmotic lysis by binding to the midgut
CC epithelial cells of mosquitoes.
CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; D88381; BAA32397.1; -; Genomic_DNA.
DR HSP; Q06117; J16.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR Pfam; PF03944; Endotoxin_C; 1.
DR Pfam; PF00555; Endotoxin_M; 1.
DR Pfam; PF03945; Endotoxin_N; 1.
KW Plasmid; Sporulation; Toxin.
SQ SEQUENCE 682 AA; 78491 MW; 5351EA63E2B042F7 CRC64;

Query Match 18.3%; Score 709; DB 1; Length 682;
Best Local Similarity 30.3%; Pred. No. 3.2e-37;
Matches 233; Conservative 112; Mismatches 277; Indels 144; Gaps 30;

QY 1 MNQNDNNEVEIIDS-----HTSPVFPNRNDSRYPVTNNPNQPLQNTYKELNWCQGN 55
DB 1 MNSYQNKVEIILDAKENTCHMSNCP-----KYPLANDPQMYLRNTHYKDMINMCEE 53

QY 56 NTOYGNFETPASADTIAAVSAGTIVSGTLGAGLTSISGPIGTGIIAISFGTLTV 115
DB 54 -----ASYASSGPSQLFKVG-----GSIVAKILGMIPEVGPL-----LSWMVSL 92

Qy	116	FWPAGEQDKTWTQBIKMGELPVDTPLPSTESIKQLKLTLEGFQILOSNTALDDWRKUK	175
Db	93	FWPTEEEKNTWEDMIKYVALLKOELTNDTILNRATSNLSGLNESLNIYNRALAAWKQNK	152
Qy	176	RLQAPGLPPSSALQQAALTLKIRPENVDNDFIREPG-FQLETYKTLPLPIYAAQHHL	234
Db	153	NNFASG-----BLIRSIINDLHILFRDIQSDPFLUGGYETVLPLPYSASANLHL	201
Qy	235	NLLQO----GAELADEWNADIHPSQIEPNAGTSDDYIKLLKENIKPYSNYCANTYRTGLK	290
Db	202	LLLRDVALYKELG-----YPSI-----DVBFYNEQKYTEKYSNYCVNTYKSGLE	248
Qy	291	NLRDEPNMKWSIFNDYRRYMITVLDTTSQPSLYDIKRYRDSIGGIEVKGIKNEULTREIY	350
Db	249	S---KKQIGWSDFNRYREMTLSVLDDIVALPFLYDTGLYPKSDGKHVKA---ELTREIY	302
Qy	351	TTEINFRLPOLRYOPNLA--TMEYNLRASPKLPSLEQRIFFYENTNPGNRLVCISNR	408
Db	303	SDVIN-DHRYGLMW-PYISFHAESLYTRP-HAFTWLKGRFVNSINSTWFLSGGENR	359
Qy	409	DAPTY-SNTI-----TETLYGERTGSPTKTIRPFESYKYSIVTDROSPVSPIQPHF	460
Db	360	YFLTHGEGTINGPPLQDTEYG-----GTSSYIDISNNSSIYNLWTXKYEWIYPWTDVP	414
Qy	461	IINQIELYLGNSNNTLYKSAGGSLSNYQNTTFFQFPRKDCNLVIDPCGSP-NFNNYSH	519
Db	415	NITKINFISITDNSSESISYGAERNKPTVRTDFNF-----LLNRAGNGPTTYNDYNH	467
Qy	520	ILSHFSL-----FYSYVIGLQLOILDTGLVGLGTHSSVDRNRNALSDKLIITMIPA	570
Db	468	ILSYMLINGETFGOKRHHYSF-----AFTHSSVDRNTYTVDPDKIVQIPAVK	513
Qy	571	GNNLDTNKSVMTEGPGHTGNNLYLOSQGRLEITCETPNSTOSYFIRLRYATNGAGNTLP	630
Db	514	TNLVGAN--IIKGPGHTGDDLKUEYERFLSURIKL-FASMTFRIRIRYASNISCOMMIN	570
Qy	631	ISLPIPGVIGIPPORLANTFSGTYNNNIQYDGF-----YFQPPSTVTPLPNNIPFI	683
Db	571	IGYQNPTYFNIIPT-----ISRDYTELKPEDFQLVDTSYIYSGGPS-----ISSNTLWL	619
Qy	684	FNRAVNSIILIIKIEPIPTSMHQNREKQKLETIOTKINTFF	728
Db	620	DN----FNGSPVIIIDKIEPIPLGITLQO---AQGYDTQDNANGMY	658

RESULT 15
Q6BE06_BACTU Q6BE06_BACTU PRELIMINARY; PRT; 829 AA.
AC Q6BE06;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Cancer cell-killing Cry protein.
GN Name=Cry41Abi;
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=1428;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=A1462;
RA Yamashita S., Saitoh H., Katayama H., Akao T., Mizuki E., Park Y.,
RA Ohta M., Ito A.;
RT "cell-killing toxin gene and other genes in 6,698bp DNA from Bacillus
RT thuringiensis.";
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB116651; BAD35163.1; -; Genomic DNA.
DR GO: GO:0005102; F:receptor binding; IEA.
DR GO: GO:0006952; P:defense response; IEA.
DR GO: GO:0009405; P:pathogenesis; IEA.
DR InterPro: IPR001178; Endotoxin.
DR InterPro: IPR005638; endotoxin C.

```

DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR000772; Ricin B lectin.
DR Pfam; PF03944; Endotoxin_C; 1.
DR Pfam; PF00555; Endotoxin_M; 1.
DR Pfam; PF03945; Endotoxin_N; 1.
DR Pfam; PF00652; Ricin B lectin; 1.
DR SMART; SM00458; Ricin; 1.
DR PROSITE; PS00231; RICIN B LECTIN; 1.
DR SC SEQUENCE 829 AA; 93841 MW; 5BFF13BEC05F97A3 CRC64;

Query Match
Best Local Similarity 18.2%; Score 705; DB 2; Length 829;
Matches 213; Conservative 113; Mismatches 276; Indels 180; Gaps 28;

Qy 1 MNQNNDNNEYIIDSHTSPYFPNRNSNDSRYPTNNPNQPLONTNYKEWLNMCQGN--TQ 58
Db 1 MNQSCNNGYEVLNS-----CKGYCQCPYPAQAPGSELQNGWYKEMWNMCTSGDPTV 53

Qy 59 YGNFETFASADYIAAVSAGTIVSGTLLAGIGLTSIS-GPIGIGAIISFGTLITVFW 117
Db 54 LGGGY-----SADVKDAVITSINTASYLL-----SVPPPPAGVAAGIL----GALLGLLW 99

Qy 118 PAGEQDKTVMTQFIKMGEIFVDTPLETYESIKOLKLOTLQLEGPROILQSYNTALDDWRKLKRL 177
Db 100 PTTNQ--AVWEAFWTVTEALINOKLDEYAKSKAISLNGLKNVLEIYQDRAADWNE---- 153

Qy 178 QAPCLPSSALQAAALTAKIRFENVHNDFTREIPGFOLEYTKTLLPIYAAQANFHNL 237
Db 154 ----NPGDLRKNRVLT---EFENVNGHPENSMPSPAVRNFEVNLPIVYAEANLHLL 206

Qy 238 QQGAELADEWADIHPSQTEPNAGTSDDYYKLLKENIPKYSNYCANTYRTGLKNLRD-EP 296
Db 207 RDAVKFGEQGMGMDTP-----CAERDDMYRLRSRTEIYTDHCVNTYNGLOQAQSLQA 260

Qy 297 NM-----KWSIFNDYRRVMTITVLDTISQFSLYDI 326
Db 261 NVSDYGRYPWTQYNGSGGSYRAKEGYRTGNWNLNAPRRDITLVLDTIAQFPITYD- 319

Qy 327 KRYRDSIGGTEVKGIKNELTREIYTTTEINFDRLPQLRVOPNLATMBYNLTR-----ASFKL 382
Db 320 -----PGLYSRPVKSELTSREYTD-----IRGTTWRSDANLNTIDAIENRMVGSRLQL 368

Qy 383 PSFLEQIFVYENT-----NFGNRLVGISNRDAPYNTITETLYGERTGSPPTKTIRP 436
Db 369 FTWLTEMKFYIRNTGSITSYTHGDLVAGLEKKIRKTNDDNQWLPLEGQNTS--YTRIDRP 426

Qy 437 -----EES--YKVSIVTDROSPPVSPIQPHFIINQIELYNGSSNNTL 477
Db 427 GIELGKNYNYARTQWFETRLQLMWNTDVLNLNAGTVGNENFARWDVDPDRIYNYARST 486

Qy 478 KYSAGGSLSNYQNTTFFQPRKKDCNLVIDPGCSPNFNPNYSHLTSLSFTYSYVIGLQL 537
Db 487 NH-----FIENHRLSWIKFEPVRDNC-----PPAWPGYKQLSALL----- 521

Qy 538 QILDGTGLVGHWSVDRYNAISDKIITMIPAIGKNLDTNSKVIEGFGHTGHNLYVLQ-- 595
Db 522 -----FGWTHNSVDPEFTFIASDRITQIPAVKGYLDVNGATVVRGPGNTGGDLVRLPAY 574

Qy 596 ----SQGRLEITCETNSQSYRILRYATNGAGNTLPNISLTIPGVIPIPPQ----- 644
Db 575 NQOWTQLRVKVRPSTTARTGYNVRIRYASEG-----NANLFPVKGYVDTPANRFYETGNY 628

Qy 645 RLNNTSEGT-NYNNLVQGD-FGYFQFPSTVTLPLNRNIPPIFNRAVDSNLSILIDKIEFI 702
Db 629 AVNQTFSGSWTNSFKYLDIAIGFAANEEEFIELRCN-----SCGPYIDKIEFI 678

Qy 703 PI 704
Db 679 PV 680

Search completed: December 4, 2005, 13:08:33
Job time : 234 secs

```

Search completed: December 4, 2005, 13:08:33
Job time : 234 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2005, 12:42:36 ; Search time 48 Seconds
(without alignments)
1265.971 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869

Sequence: 1 MNQNNNVEYIIDSHTSPY.....KLETIQKINTFFTHKTL 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5 COMB.pep:*

2: /cgn2_6/ptodata/1/iaa/6 COMB.pep:*

3: /cgn2_6/ptodata/1/iaa/H COMB.pep:*

4: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep:*

5: /cgn2_6/ptodata/1/iaa/RE COMB.pep:*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1060.5	27.4	1180	2	US-09-224-024-28
2	1060.5	27.4	1180	4	PCT-US94-07902-28
3	895.5	23.1	686	2	US-09-224-024-31
4	895.5	23.1	686	4	PCT-US94-07902-31
5	775.5	20.0	1169	1	US-08-315-468-4
6	676	17.5	651	1	US-08-315-468-6
7	676	17.5	651	2	US-07-941-650A-4
8	676	17.5	659	2	US-08-996-441B-112
9	676	17.5	659	2	US-08-993-172A-112
10	676	17.5	659	2	US-08-993-170A-112
11	676	17.5	659	2	US-08-993-775B-112
12	676	17.5	659	2	US-09-427-770-112
13	676	17.5	659	2	US-09-427-769-112
14	663	17.1	652	2	US-08-996-441B-38
15	663	17.1	652	2	US-08-993-722A-38
16	663	17.1	652	2	US-08-993-170A-38
17	663	17.1	652	2	US-08-993-170A-38
18	663	17.1	652	2	US-08-993-775B-38
19	663	17.1	652	2	US-09-427-770-38
20	660	17.1	652	2	US-08-996-441B-110
21	660	17.1	652	2	US-08-993-722A-110
22	660	17.1	652	2	US-08-993-170A-110
23	660	17.1	652	2	US-08-993-170A-110
24	660	17.1	652	2	US-09-377-466B-4
25	660	17.1	652	2	US-09-427-770-110
26	660	17.1	652	2	US-09-427-769-110
27	660	17.1	652	2	US-10-232-665-4

28	660	17.1	652	4	PCT-US92-00040-2	Sequence 2, Appli
29	658	17.0	652	2	US-08-996-441B-28	Sequence 28, Appl
30	658	17.0	652	2	US-08-996-441B-36	Sequence 36, Appl
31	658	17.0	652	2	US-08-993-722A-28	Sequence 28, Appl
32	658	17.0	652	2	US-08-993-722A-36	Sequence 36, Appl
33	658	17.0	652	2	US-08-993-170A-28	Sequence 28, Appl
34	658	17.0	652	2	US-08-993-170A-36	Sequence 36, Appl
35	658	17.0	652	2	US-08-993-775B-28	Sequence 28, Appl
36	658	17.0	652	2	US-08-993-775B-36	Sequence 36, Appl
37	658	17.0	652	2	US-09-427-770-28	Sequence 28, Appl
38	658	17.0	652	2	US-09-427-770-36	Sequence 36, Appl
39	658	17.0	652	2	US-09-427-769-28	Sequence 28, Appl
40	658	17.0	652	2	US-09-427-769-36	Sequence 36, Appl
41	656	17.0	652	2	US-09-377-466B-6	Sequence 6, Appli
42	656	17.0	652	2	US-10-232-665-6	Sequence 6, Appli
43	655	16.9	652	2	US-08-996-441B-18	Sequence 18, Appl
44	655	16.9	652	2	US-08-996-441B-60	Sequence 60, Appl
45	655	16.9	652	2	US-08-993-722A-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-09-224-024-28
; Sequence 28, Application US/09224024
; Patent No. 6056953
; GENERAL INFORMATION:
; APPLICANT: Leslie Hickie
; APPLICANT: Jewel Payne
; TITLE OF INVENTION: Materials and Methods for the Control of
; TITLE OF INVENTION: Calliphoridae Pests
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/224,024
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/856,226
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA79
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1180 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-224-024-28

Query Match 27.4%; Score 1060.5; DB 2; Length 1180;
Best Local Similarity 35.1%; Pred No. 1.2e-89;
Matches 271; Conservative 117; Mismatches 279; Indels 105; Gaps 24;
Oy 1 MNQNNNVEYIIDSHTSPYFPNRSNDSRYPTNNPNQPLQNTYKWLNNCOGNTQYG 60


```
; REFERENCE/DOCKET NUMBER: MA79
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 686 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-07902-31

Query Match      23.1%; Score 895.5; DB 4; Length 686;
Best Local Similarity 31.9%; Pred. No. 1.4e-74;
Matches 245; Conservative 127; Mismatches 267; Indels 129; Gaps 30;

QY 1 MNQNNNEVEIIDSHTSPYFPNNSNDSPYTNPNQPLQNTYKWLNMCOGNTQYG 60
DB 1 MNPYQNKNEYIPNAPNGF--SKSNYSRYPLANKPNQPLKNTYKDWLNVCODNQYQ 58
QY 61 DNFETPASADTTAAVSAGTIVSGTLLAGIGLTSISGPIGIIIGAIISFGTLITVFWPAG 120
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 109
59 NNAGNFVSSETIVGVSAGIIVGTMLG-----APAAP--VLAAGIISFGTILLPIFW-QG
QY 121 EODKTWTOFIKNGEIFVDTPLTE---SIKQLKLTLEGFRQILQSYNTALDDWRKLR 177
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 163
110 SDPANWQDLNLIG---GRPIQIDKNIINVLTSIVTPIKNLDKYQEFDFKWEPA-- 163
QY 178 QAPGLPPSALQQAALTILKIRFENVN---DFREIPGFOLEYTKYLLLIPIYAQAANFHL 234
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 211
164 -----THANAKAVHDLFTLEPIIDKOLDMLKNNASYRIPT-----LPAYAQIATWHL
QY 235 NLLQQAELADENADIHSPQIBNAGTSDYK-LKKNIPKYSNYCANTYTGKLNLR 293
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 268
212 NLLKHAATYNIW---LQNGINPSTNSNYQGLKRIQBYTDYCIOTYNAGLTMIR
QY 294 DBPNMKWSIFNDYRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKRKNELTREIYTT 353
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 319
269 TTNATWNYNTYRLWTLTLVLDLIAFPNYDEKY-----PIGVKSELREYV-TN
QY 354 INFDRLPQURVQPNLATMEYNLTRASFKLPSFLQEIFYFNTN-----FGNRL 402
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 372
320 VNSDTP-----RTITELENGLTR-NPTLFTWINGRFRYTRNSRDILDPYDIFSFTGNQM
QY 403 VGISNRDAPYSTNIETLYGERTG---SPTTKTIRPFESY----KVISIVTRQSPVPSP 455
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 424
373 -----AFTHNDDRNIIWGAHVGHIIISQTSKVPFPYRNKPIDKVEIVRHREYSDI-
QY 456 IQPHFIINOIELYINGSNNTLYKSAGGSL-SNYQNTTFFQPRKXCDNLVIDPGCSPNF 514
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 464
425 -----YEMIFFNSSEVFYSNSTIENNYKRTDSYMIKQTKWN-----
QY 515 NNYSHILSHFLPTYSYVIGLQILDTGVLGWTHSSVDRYNAISDKIIMPIAKGNL 574
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 521
465 KEYGHTLSYKTDNYIFSVVRRRRV---AFSWTHTSVDQNTIDLDNIQIHALKALKV
QY 575 DTNSKVIKPGHGTGNLVLOSQRLEITCETNSQSYFIRLRYATNGAGNTLPIISLT 634
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 575
522 SSDSKIVKPGHGTGGDLVLKDSMDFRVRP-LKNVSRQYQVRIYATNA-----PKTTVF
QY 635 IPGVIGIPQRLNNTPSGTNYN--NLQYDGFYGFQFPSTV-----TLPNLRNIPFI 683
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 629
576 LTGIDTISVE-LPSTTSRQPNATDITYADFGVYTPRTVPNTKTFEGEDTLMT-----L
QY 684 FNRADVSNILIDKIEPIPTSMHQNRKQKLETIQTKINTFFTNH 731
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
630 YGTPNHSYNI-YIDKIEFIPITQSVLDYTEKQNIKTQKIVNDLFVNN 676

RESULT 5
US-08-315-468-4
; Sequence 4, Application US/08315468
```

```
; Patent No. 5554534
; GENERAL INFORMATION:
; APPLICANT: Michaels, Tracy Ellis
; APPLICANT: Poncerrada, Luis
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Process for Controlling Scarab Pests
; TITLE OF INVENTION: with Bacillus thuringiensis Isolates
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/315,468
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,941
; FILING DATE: 01 FEB 1993
; APPLICATION NUMBER: 07/828,430
; FILING DATE: 30-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/808,316
; FILING DATE: 16-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA73.C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1169 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: kumamotoensis
; INDIVIDUAL ISOLATE: 50C
; IMMEDIATE SOURCE:
; LIBRARY: LambdaGEM-11 library of L. Poncerrada
; CLONE: 50C(b)
US-08-315-468-4

Query Match      20.0%; Score 775.5; DB 1; Length 1169;
Best Local Similarity 28.5%; Pred. No. 6.4e-63;
Matches 217; Conservative 136; Mismatches 289; Indels 119; Gaps 27;

QY 5 NDNNEVEIIDSHTSPYFPNNSNDSPYTNPNQPLQNTYKWLNMCOGN-TQYGDN 62
DB 4 NNQNEVEIIDSHTSPYFPNNSNDSPYTNPNQPLQNTYKWLNMCOGN-TQYGDN 59
QY 63 FETPASADTTAAVSAGTIVSGTLLAGIGLTSISGPIGIIIGAIISFGTLITVFWPAG 122
DB 60 PEVFLSEQD--AVKAAIDIVGKLITGLG-----VPFVGPIVSLYVQLDILWPS--K 107
QY 123 DKTWTOFIKNGEIFVDTPLTESIKQLKLTLEGFRQILQSYNTALDDWRKLRQAPGL 182
DB 108 QKSQWEIFMQVEELINOKIAEYARNKALSEGLEGNNGNYQLYALAEWKE----- 158
```

```
QY 183 PPSSALQQAALTLKIRFENVHNDIFIREIPGFOLEYTKTLPLPIVAAANFHLNLQOGAE 242
Db 159 NPNGS--RALDVRNRREILDSLTQYMPSRVNFVFPFVTVYMAANLHLLLRDASI 216
QY 243 LADBNADIHPSQIEPNAGTDDYYKLLKENIPKYSNYCANTYRTGLKNLRDEPNMKWSI 302
Db 217 FGEWGL-----STSTINNYNRQMKLTAEYSDHCVKWTETGLAKLGSSAKQWID 267
QY 303 FNDYRRYMTITVLDTISQFSLYDIKRYRDSGGIEVGKIKNELREIYTTINFDRLPOL 362
Db 268 YNOFRREMTITVLVDVALFSNYDRTY-----PLATTAQLTREVVYDPLGAVDVENI 319
QY 363 RVQPNLATMEYNLTRASF--LFSLEQOFIYTTENTNF-----GNRLVGISNRD 409
Db 320 GSWYDKAPSESEIEKAIRPHVDYITGLTVYTKRSFTSDRYMYWAGHQ---ISYKH 376
QY 410 APTYSNTITELYGERTGSPTKTIRPFESYK-----SIVTDROSPVVS-----PI 456
Db 377 IGT--SSFTQ--MYGTNQLQSTSNF-DFTNYDIYKTLNSGAVLLDIVPGYTYTFFGMP 433
QY 457 QPHFIINQIELYNGSSNNTLYKSAGGSLSYQNTFFQPPRKDCNLIIDPGCS--PNF 514
Db 434 TEFPMVNL-----NTRK-----TLTYKPAKDIIDTRDSELELPETSGQPNY 479
QY 515 NNYSHILSH-----PSLFTYSYVIGLQILDGTGLGTHSSVDRYNAISDKIITMIPA 570
Db 480 ESYSHRUGHITFYSSSTSTV-----PVFSWTHRSADLTTNTVKSGBEITQIPGGK 529
QY 571 GNNLDTNSKVIEGHGTGNNLVYLSQ--GRLEITCETPNSTQSYFIRLRYATNGAGNLP 629
Db 530 SSTIGRNTYIYKRGYTGGLVALTDRIKSGCFQWIPES--QRFIRIRYASNETS-----584
QY 630 NISLTIGVIGIPQRLNNTFSGYNNYNNLVQDGFYGFQFPFTVTLPLNRNIPFIFNRADV 689
Db 585 --VISYGLAQSGTLKFNQYTSKNENDLTYNDPKYIEYPRVISWASSNITQLRSIGT 642
QY 690 SNSLITDKIEFTPIITSSMHQNRKQKLETTQTKINTFFTN 730
Db 643 NTNLFILDRIEFIPVDETYEATD---LEAKKAVNALFTN 680

RESULT 6
US-08-315-468-6
; Sequence 6, Application US/08315468
; Patent No. 5554534
; GENERAL INFORMATION:
; APPLICANT: Michaels, Tracy Ellis
; APPLICANT: Fonceerrada, Luis
; TITLE OF INVENTION: Process for Controlling Scarab Pests
; TITLE OF INVENTION: with Bacillus thuringiensis Isolates
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/315,468
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,941
; FILING DATE: 01 FEB 1993
; APPLICATION NUMBER: 07/828,430
; FILING DATE: 30-JAN-1992
```

```
;
; PRIORITY DATA:
; APPLICATION NUMBER: 07/808,316
; FILING DATE: 16-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA73.C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 651 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: tolworthi
; INDIVIDUAL ISOLATE: 43F
; IMMEDIATE SOURCE:
; CLONE: E. coli XL1-Blue (pMI,98-4), NRRL B-18291
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..651
; US-08-315-468-6
```

Query Match 17.5%; Score 676; DB 1; Length 651;

Best Local Similarity 28.0%; Pred. No. 5e-54;

Matches 208; Conservative 116; Mismatches 286; Indels 134; Gaps 28;

```
QY 1 MNQNDNNEYIIDSHTSPYFNPNRNSDSRYVYTNPNQPLQNTYKELNMCQGTQYG 60
Db 1 MNPNN-RSEYDTIKVTPNSELP---TNHNPPLADNPSTLBEINKEFLMTADNST--54
QY 61 DNFETASADTAAVSACTIVSGTLLAGIGLTSISGPIIGIIGAIISFGTLITVFWPAG 120
Db 55 ---EVLDSSTVKDAVGTGIVVVGQILGVVG-----VPPFAGALTSFYQSFLNALWPS-102
QY 121 EQDKTVTQPIKMGHIFVDTPLTESIKQLKLTLEGFROILOSINTALDDWRKLRLOAP 180
Db 103 --DADPWKAFMAQVEVLIDKKIEEYAKSKALAELOGLQNFEVDYNALDSWKK-----153
QY 181 GLPPSSALQQAALTLKIRFENVHNDIFIREIPGFOLEYTKTLPLPIVAAANFHLNLQOG 240
Db 154 -APVNLRSRRSQDRIRELFSQAESHFRNSMPSFAVSKFEVLFLPTTYAQAAHTHLLLKDA 212
QY 241 AELADEMNADIHPSQIEPNAGTSD--YYKLLKENIPKYSNYCANTYRTGLKNLRDEPN 297
Db 213 QVGEWGY-----SSEIAEFYQRLKLTQOYTDHCVNWNVYVGLNSLRGSY 260
QY 298 MKWSIFNDYRRYMTITVLDTISQFSLYDIKRYRDSIGGIEVGKIKNELREIYTTINFD 357
Db 261 DAWKFNRRFRREMTITVLVDLIVLPPFYDRLYS-----KGVKTELTRDIFTDPI-FT 311
QY 358 RLPQLRVQPNLATMEYNLTRASF--LFSLEQOFIYTTENTNFNRL-VGISNRDAPTY---413
Db 312 LNALQYGFPTFSIENSIRKP--HLFDYLRIGIEFT-----RLRPGYSGKDSFNYWSG 362
QY 414 -----SNTITELYGERTGSPTKTIRPFES---YKVSIVTDROSPVVSIPQH 459
Db 363 NYVETRPSIGSNDITSPFYGDKSTIEPQKL--SFDGQKVYRTIANTDIAAPDGKI--Y 418
QY 460 FIINQIEL--YLANGSSNNTLYKSAGGSLSYQNTFFQPPRKDCNLIIDPGCS--PNFNN 516
Db 419 FGVTKVDFSQYDDQKNET-----STQYDSKRYNGYLGADSIDQLPPTTDEPLEKA 471
QY 517 YSHILSHFSLTYSYVIGLQILDGTGLGTHSSVDRYNAISDKIITMIPA KGNLDT 576
Db 472 YSHQLNVAECFL-----MQDRRGITPFFTWTTHRSVDFTIDAEKITQLPVVYKAYALSS 525
```



```

; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996.441B
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: MECO:151
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-996-441B-112

Query Match 17.5%; Score 676; DB 2; Length 659;
Best Local Similarity 28.0%; Pred. No. 5,1e-54;
Matches 208; Conservative 116; Mismatches 286; Indels 134; Gaps 28

Qy 1 MNQNDNNEYEIIISHTSPYPFNRNSNDSRYPYTNPNQPLQNTNYKEWLNMCQGNTOYQ 60
Db 9 MNPNN-RSEYDIKVTNPSELP---TNNQYPLADNPNSTLEELNYKEFLRMTADNST-- 62
Qy 61 DNFTFASADTIAVSACTIVSGTLLAGIGLTSISGPIGIGAILISFGTLITVFWFAG 120
Db 63 ---EVLDSSTVKDAGVTGISVVGQILGVVG-----VPPAGALTSPYQSFNLAINWFS- 110
Qy 121 EQDKTVMVTFQIKMGIEIFVDTPLTETSIKQLQTLQEGFQIILQSNVNTALDDWRKLRLOAP 180
Db 111 --DADPKAFNAQVLEVLIDKKIEEYAKSAEALQGLQNNPEDIYNALDSWK----- 161
Qy 181 GLPSSALQQAALTKIRFNVNHDFTREIRPGQLETYKTLTLLPTIYAQAAPHMLMLQOG 240
Db 162 -APVNLRRSRQSDRIRELFSQAESHFRNSMPFSFAVSKPEVLPLPTYAQAANTHLLLKDA 220
Qy 241 AELADENWADIHPQIEENACTSD--YYKLLKENIPKYSNYCANTYRTGLKNLRDPE 297
Db 221 QVFGEEGY-----SSEDAEFYQRLKLTQYTDHCNVNWNVLNLSRGSTY 268
Qy 298 MKWEIFNDYRRYMTITVLDTISQCSLYDIKYRDSIGGIEVKGIKNELTREITYTEINFD 357
Db 269 DAWKFNRFREMITLVDLLVLFPFYDVRYS-----KGVTETLRDITFDPI-FT 319
Qy 358 RLPOLRVQPNLATMEYNLTRASPKLFSLEOFIFYTENTNFGNRL-VGINSRDAPTY--- 413
Db 320 LNALQYEGPTFSIENSIRKP-HLFDYLRGIEFHT-----RLRPGVSGKDSFNYWSG 370
Qy 414 -----SNVITETLYGERTGSPPTKIRPPES---YKVSIVTRQSPPVSPIDPH 459
Db 371 NYVETRFPSIGSNDRITTSFPYGDKSIEPIQKL--SPDGQKVRRTIANTDIAAPPDGI--Y 426
Qy 460 FIINQIEL-YLNGSSNNTLKYTSAGGSLSNYQNTIFFQPRKKOCNVLDPGCS--PNFNN 516
Db 427 FGVTKVDFSQVDDQKNET-----STQYTSKRYNGYLGQAQSDIDQUPPETTDELEKA 479
Qy 517 YSHLSHFSFTYSVIGLOQLDITGLVGLTHSSVDRYNAISDKIIITWIPAIKGNNDLT 576

```

Db 63 ----EVLDSSTVDAVGTGIVGQILGWG-----VFPAGALTSFYQSFNLNMFPS- 110
 Qy 121 EODKTVWTFPIKGEIFVDTPLTESIKQLKLTLEGFRQLQSNTALDWRKLRLOAP 180
 Db 111 --DADPKAFMAQVEVLIDKKIEYAKSKALAEQLQNNFEDVYVALDSWK----- 161
 Qy 181 GLPPSSALQQAULTKIRFENVHNDIFREIPGQLETYKTLPLPIYAQAANFHLNLOOG 240
 Db 162 -APVNLRSRQRIREFLSQAESHFRNSMPSPAVSKFEVLFLPTYAQAANTHLLLKDA 220
 Qy 241 AELADEWNADIHSQIEPNAGTSD--YYKLKENIPKYSNCANTYRGLKNLRDEPN 297
 Db 221 QVFEWGY-----SSEDIAEFYQRLKLTQYTDHCNVNMYNGLNSLRGTY 268
 Qy 298 MKWSIFNDYRYMTITVLDTISQSLYDIKRYRDSIGGIEVKGKIELTREIYVTEINF 357
 Db 269 DAWKFNRRFRMTLVLDLIVLPFYDVRYS-----KGKTELTRDIFTDPI-FT 319
 Qy 358 RLPQLRVQPNLATMEYNLTRASPKLFSLEQFIYFTENTNFGNRL-VGISNRDAPTY--- 413
 Db 320 LNALQEVGPTFSSIRKP--HLFDYLRGIEFHT-----RLRPGYSGKDSFNYSWG 370
 Qy 414 -----SMTITELTGERTGPTTKTIRPFES---YKSVITDRQSPVPSPIQPH 459
 Db 371 NYVETREPSIGSNDTITSPFYGDKSIEPIQL--SFDQKQVYRTIANTDIAAFPDKI--Y 426
 Qy 460 FIINQIEL-YLNGSSNNLTKYSAGGSLSNQTNTFFQPKKCNLVDPGCS--PNFNN 516
 Db 427 FGVTKVDFOYDQKNET-----STQYDSKRYNGYLGAQDSIDQLPETTDEPLEKA 479
 Qy 517 YSHILSHFSLFYVYVIGLQILDTGLVGTWTHSSVDYRNAISDKIITMIPAIGNNLDT 576
 Db 480 YSHQLNVAECFL-----MDRRGTIPFTTWTHRSVDFNTDAEKITQLPVVKAYALS 533
 Qy 577 NSKVIIEGPGHTGNLVYLO-----SQGRLEITCTPNSTQSYFIRLRVATNGAGNTLPNIS 632
 Db 534 GASIIIEGPGTGNLLFLKSSNSIAKFKVTLNSAALLQRYRIRYAST-----TNLR 587
 Qy 633 LITPGVIGIPPORLNTFSGTNNYLNQY--GDFGY--FQFPSTVTLPLNRPINR 688
 Db 588 LfV-----QNSNDFLVIYINKTNWIDGDLTYQTFDFATS-----NSNMGFSGDTND 634
 Qy 689 -----VSNLIIIDKIEFPI 704
 Db 635 FIIGAESFVSNEKIYIDKIEFIV 658

RESULT 10

US-08-993-170A-112
 ; Sequence 112, Application US/08993170A
 ; Patent No. 6063597
 ; GENERAL INFORMATION:
 ; APPLICANT: English, Leigh H.
 ; APPLICANT: Brussock, Susan M.
 ; APPLICANT: Malvar, Thomas M.
 ; APPLICANT: Bryson, James W.
 ; APPLICANT: Kulesza, Caroline A.
 ; APPLICANT: Walters, Frederick S.
 ; APPLICANT: Slatin, Stephen L.
 ; APPLICANT: Von Tersch, Michael A.
 ; TITLE OF INVENTION: POLYPEPTIDE COMPOSITIONS TOXIC TO
 ; COLLEPTERAN INSECTS
 ; NUMBER OF SEQUENCES: 113
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Arnold, White & Durkee
 ; STREET: P.O. Box 4433
 ; CITY: Houston
 ; STATE: Texas
 ; COUNTRY: USA
 ; ZIP: 77210
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/993,170A
 ; FILING DATE: 18-DEC-1997
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kitchell, Barbara S.
 ; REGISTRATION NUMBER: 33,928
 ; REFERENCE/DOCKET NUMBER: MECO:002
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 512/418-3000
 ; TELEFAX: 512/474-7577
 ; INFORMATION FOR SEQ ID NO: 112:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 659 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; US-08-993-170A-112

Query Match 17.5%; Score 676; DB 2; Length 659;
 Best Local Similarity 28.0%; Pred. No. 5,1e-54;
 Matches 208; Conservative 116; Mismatches 286; Indels 134; Gaps 28;

Qy 1 MNQNDNNEYIIDSHTSPYFNNRNSNDSRYPTNNPQNTNTYKEWLNMCQNTQYG 60
 Db 9 MNPN--RSEYDTTKVTENSELP---TNHQYPLADNPNSTLEELNYKEFLRMTADNST-- 62
 Qy 61 DNFETASADTIAAVSAGTIVSGTLLAGIGLTSISGPIGIIISFGLTITVFWPAG 120
 Db 63 ---EVLDSSTVDAVGTGIVGQILGWG-----VFPAGALTSFYQSFNLNMFPS- 110
 Qy 121 EODKTVWTFPIKGEIFVDTPLTESIKQLKLTLEGFRQLQSNTALDWRKLRLOAP 180
 Db 111 --DADPKAFMAQVEVLIDKKIEYAKSKALAEQLQNNFEDVYVALDSWK----- 161
 Qy 181 GLPPSSALQQAULTKIRFENVHNDIFREIPGQLETYKTLPLPIYAQAANFHLNLOOG 240
 Db 162 -APVNLRSRQRIREFLSQAESHFRNSMPSPAVSKFEVLFLPTYAQAANTHLLLKDA 220
 Qy 241 AELADEWNADIHSQIEPNAGTSD--YYKLKENIPKYSNCANTYRGLKNLRDEPN 297
 Db 221 QVFEWGY-----SSEDIAEFYQRLKLTQYTDHCNVNMYNGLNSLRGTY 268
 Qy 298 MKWSIFNDYRYMTITVLDTISQSLYDIKRYRDSIGGIEVKGKIELTREIYVTEINF 357
 Db 269 DAWKFNRRFRMTLVLDLIVLPFYDVRYS-----KGKTELTRDIFTDPI-FT 319
 Qy 358 RLPQLRVQPNLATMEYNLTRASPKLFSLEQFIYFTENTNFGNRL-VGISNRDAPTY--- 413
 Db 320 LNALQEVGPTFSSIRKP--HLFDYLRGIEFHT-----RLRPGYSGKDSFNYSWG 370
 Qy 414 -----SMTITELTGERTGPTTKTIRPFES---YKSVITDRQSPVPSPIQPH 459
 Db 371 NYVETREPSIGSNDTITSPFYGDKSIEPIQL--SFDQKQVYRTIANTDIAAFPDKI--Y 426
 Qy 460 FIINQIEL-YLNGSSNNLTKYSAGGSLSNQTNTFFQPKKCNLVDPGCS--PNFNN 516
 Db 427 FGVTKVDFOYDQKNET-----STQYDSKRYNGYLGAQDSIDQLPETTDEPLEKA 479
 Qy 517 YSHILSHFSLFYVYVIGLQILDTGLVGTWTHSSVDYRNAISDKIITMIPAIGNNLDT 576
 Db 480 YSHQLNVAECFL-----MDRRGTIPFTTWTHRSVDFNTDAEKITQLPVVKAYALS 533
 Qy 577 NSKVIIEGPGHTGNLVYLO-----SQGRLEITCTPNSTQSYFIRLRVATNGAGNTLPNIS 632
 Db 534 GASIIIEGPGTGNLLFLKSSNSIAKFKVTLNSAALLQRYRIRYAST-----TNLR 587
 Qy 633 LITPGVIGIPPORLNTFSGTNNYLNQY--GDFGY--FQFPSTVTLPLNRPINR 688
 Db 588 LfV-----QNSNDFLVIYINKTNWIDGDLTYQTFDFATS-----NSNMGFSGDTND 634

Qy 689 -----VNSILIIDKIEPIPI 704
Db 635 FIIGAESFVSNKIIDKIEPIV 658

RESULT 11

US-08-993-775B-112
; Sequence 112, Application US/08993775B
; Patent No. 6077824
; GENERAL INFORMATION:
; APPLICANT: English, Leigh H.
; APPLICANT: Brussock, Susan M.
; APPLICANT: Malvar, Thomas W.
; APPLICANT: Bryson, James W.
; APPLICANT: Kulesza, Caroline A.
; APPLICANT: Walters, Frederick S.
; APPLICANT: Slatin, Stephen L.
; APPLICANT: Von Tersch, Michael A.
; TITLE OF INVENTION: METHODS FOR IMPROVING THE ACTIVITY OF
; DELTA-ENDOTOXINS AGAINST INSECT PESTS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,775B
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: MECO:150
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear

US-08-993-775B-112

Query Match 17.5%; Score 676; DB 2; Length 659;
Best Local Similarity 28.0%; Pred. No. 5.1e-54;
Matches 208; Conservative 116; Mismatches 286; Indels 134; Gaps 26;
Qy 1 MNQNDNNEVEIDSHTSPPYFNPNSNDSRYPTNNPQPLQNTNYKEWLNMCQNTQYG 60
Db 9 MNFNN-RSEYDTIKVTNPSELPL---THNQYPLADNPNSLLEUNYKEFLRWMTADNST-- 62
Qy 61 DNFETFAADTIAAASAGTTVSGTLGAGIGGLTSISGPIGIGAIISFGTLITVFWPAG 120
Db 63 ---EVLDSSTVKDAGVGISVVGQILGVG-----VPPAGALTSPYQSFLNAIWPS- 110
Qy 121 EQDKTWTQIKGEIFVDPPLTPTESIKOLKLTLEGFRQLQSNVNTALDWRKUKRLQAP 180
Db 111 --DADPWKPAQVENVLIDKIKIEYAKAKALAEQLQNNFEDVYVALDSWK----- 161
Qy 181 GLPSSALQAAULTKIRFENVHNDFTREIPGQLEYTKLLLPYQAANFHLNLAQOG 240
Db 162 -APVNLRRSQDRIRELFSQAESHFRNSMPFSAVSKFEVFLFLPTYAQAANTHLLLKDA 220

Qy 241 AELADEWNADIHPSQIEPNAGTSDD---YYKLLKENIKPKSYNYCANTYTGKLNLRDEPN 297
Db 221 QVFGENGWY-----SSEDIAEFYQRLKLTQOYTDHCVMNWNVGLNSRGSTY 268
Qy 298 MKWSIFNDYRRYMTITVLDTISQFSLYDIKRVDSIGGIEVGKIGKIELTREIYTTTEINF 357
Db 269 DAWVKFNRFRREMTLTVLDLIVLFFPYDVRLYS-----KGVKTELTRDIFTDFI-FT 319
Qy 358 RLPQLRVQPNLATMEYNLTRASFKLPSLEQFIFYTENTNFQNL-VGISNRDAPTY--- 413
Db 320 LNALQEGYGTPESSISIRKP-HLFDVLRGIEPHT-----RLRPGYSGKDSFNWWSG 370
Qy 414 -----SNTITETLYGERTSGPTTKTIRPES-----YKVSIVTDQSPVSPIQPH 459
Db 371 NYVETRPISGNSDITITSPFYGDKSIEPIQKL--SFDGQKVYRTIANTIAAPDGGKI--Y 426
Qy 460 FIINQIEL-YLNGSSNNTLKYSGAGSLSNYQNTTFQFPRKKDCNLVIDPGCS--PNFNN 516
Db 427 FGVTKVDFPSQYDDQKNET-----STQYDSKRYNGYLGADSIDQLPETTDEPLEKA 479
Qy 517 YSHLSHPSLFYTYVYIGLQLQILDGTGLVHTHSSVDRYNAISDKIITMIPAIGNNLDT 576
Db 480 YSHQLNVAECFL-----MQDRRGTIPTFTWTHRSVDFFENTIDAEKITQLPVVKAVALSS 533
Qy 577 NSKVIIEGPGHGTGMLVYLQ-----SQRLIEITCETPNSQSYFIRLRYATNGAGNTLPNIS 632
Db 534 GASIIIEGPGFTGGNLLFLKSSNSIAKFKVTLNSAALLQRYVRIRYAST-----TNLR 587
Qy 633 LTIPGVIGIPQRLNNTSPGTYNNILQY--GDFGY--FQFPSTVTLPLNRNIPFIFNRAD 688
Db 588 LFV-----QNSNDFLVIYINKTMIDGLTYQTFDFATS-----NSNMGFSGDTND 634
Qy 689 -----VNSILIIDKIEPIPI 704
Db 635 FIIGAESFVSNKIIDKIEPIV 658

RESULT 12

US-09-427-770-112
; Sequence 112, Application US/09427770
; Patent No. 6620988
; GENERAL INFORMATION:
; APPLICANT: English, Leigh H.
; APPLICANT: Brussock, Susan M.
; APPLICANT: Malvar, Thomas W.
; APPLICANT: Bryson, James W.
; APPLICANT: Kulesza, Caroline A.
; APPLICANT: Walters, Frederick S.
; APPLICANT: Slatin, Stephen L.
; APPLICANT: Romano, Charles
; APPLICANT: Von Tersch, Michael A.
; TITLE OF INVENTION: NUCLEIC ACID SEGMENTS ENCODING MODIFIED
; COLEOPTERAN-TOXIC CRYSTAL PROTEINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/427,770
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/993,722
; FILING DATE: 18-DEC-1997

```

; ATTORNEY/AGENT INFORMATION:
;   NAME: Kitchell, Barbara S.
;   REGISTRATION NUMBER: 33,928
;   REFERENCE/DOCKET NUMBER: MECO:149
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 512/418-3106
;   TELEFAX: 512/474-7577
;   INFORMATION FOR SEQ ID NO: 112:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 659 amino acids
;       TYPE: amino acid
;       STRANDEDNESS:
;       TOPOLOGY: linear
; US-09-427-770-112

Query Match      17.5%; Score 676; DB 2; Length 659;
Best Local Similarity 28.08; Pred. No. 5.1e-54;
Matches 208; Conservative 116; Mismatches 286; Indels 134; Gaps 28;

QY 1 MNONDNNEVEIIDSHTSPYPPNPNNSNDSRYPTNNPNQPLQNTNYKEWLNMCGNTQYG 60
Db   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
9 MNPNN-RSEYDTIKVTNPSELP---TNHNGPLADNPNSLTLEELNKEYEFLRWLTADNST-- 62
QY 61 DNPETFASADTIAAVSAGTIVSGTLLAGILGGLTISIGPIGIGAILIISFGTLIIVFPAG 120
Db   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
63 ---EVLDSSTVKDAGVGTISVVGQILGWG-----VPPAGALTSFYOSFLNAIWPS- 110
QY 121 EQQKTVMTQFIKMGIEIFVDTPLETSIKOLKIQLEGFQRILQSYNTALDDWRKLRLOAP 180
Db   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
111 --DADPKWPAQVEVLIDKKIEYAKSALAEQLQNNFEDYVNALDSWK----- 161
QY 181 GLPSSALQQAALTIKIRFENVHNDFIREFPGOLETYKTLTLLPIYAQAANPHNLQOQ 240
Db   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
162 -APVNLRSRSDQRIREFLSQAESHFNPSMPFAVSKFEVFLFLPTYAQAANTHLLLKDA 220
QY 241 AELADENADIHSPQIEPNAGTSD---YKLLKENIPKYSNYCANTYRTGLKNLRDEPN 297
Db   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
221 QVFGEWGY-----SSEDIAEYQRLKLTQOYTDHCNVNMYNVLGNSLRGSTY 268
QY 298 MKWSIFNDYRYMTITVLDTISQPSLDYIKRYRDSIGIEVKGIKNEITREIYVTEINFD 357
Db   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
269 DAWKFNRFREMTLTVDLILVLPFYDVRYS-----KGVKTELRDIFTDPI-FT 319
QY 358 RLPQLRVQPNLATMEYNLATRSKFLBFLBQFIYTTENTNFGNRL-VGISNRDAPTY--- 413
Db   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
320 LNALQEXGPTFSIENSIRKP--HLFDYLRGIEFT-----RLRPGYSGKDSFNWSG 370
QY 414 -----SNTIETLYCERGTSPTTKTIRFPES---YKVSIVTDRQSPVPSPIQPH 459
Db   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
371 NYVETRPSIGSNTDITSPFYGDKSIEPIQL--SFDGQKVYRTIANTDIAAFPDKGI--Y 426
QY 460 FIINQIEL-YLNGSSNNNTLKYSGAGSLSNTQNTTFQPRKCKNCLVIDPGCS--PNFNN 516
Db   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
427 FGVTKVDFSDYDQKNET-----STQTVDSKRYNGYLGQDSIDQIPETTIDPELEKA 479
QY 517 YSHLSHFSLFTYSYVIGLOQLDITGVLGWTHSSVDRYNAISDKIITMIPAIGKNLDT 576
Db   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
480 YSHQLNVAECFL-----MQDRRGITPEFTWTWTHRSVDFTIDAEEKITQLPVVKAYALS 533
QY 577 NSKVIIEGPGHTGNLVLO-----SQGLEITCETPNSTOSYFIRLRVATNGAGNTLPNIS 632
Db   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
534 GASIIIEGPGTGNLFLPKSSNSIAKFKVTLNSAALLQRYRVIRYAST-----TNLR 587
QY 633 LTIPGVIGIPPPQRLNNTFSGTNTVNNLQY--GDFGY--FQPPSTVTLPLNRNIPFIENRAD 688
Db   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
588 LFV-----QNSNDFLIYINKTNWIDGLTYQTDFATS-----NSNWGFSGTND 634
QY 689 -----VNSNLIIDKIEFPI 704
Db 635 FIICGASFVSNKIIYIDKIEFIPV 658

```

QY 298 MKWSIENDRYRYMTITVLDTISQFSLYDIKEVRDSIGGIEVKGIGKVELTRIIYITEINF 357
 Db 269 DAWVKFNRFRREMTLTVLDLVLVLPFFVDRLYS-----KGKVELTRDIFTDPI-FT 319
 QY 358 RLPLQRVQPNLATMEYNLTASFLFSLFQFIFVTENTNFGNRL-VGISNRDAPTY--- 413
 Db 320 LNALQEVGPTFSSISIRKP--HLFDYLRGIEFHT-----RLRPGYSGKDSFNYWSG 370
 QY 414 -----SNTITETLYGERTGSPPTKTIRPPES---YKVSIVTDROQSPVPSPIQPH 459
 Db 371 NYVETRPSIGSNDTITSPFYGDKSIEPIQL--SPDGQKVYRTIANTDIAAPPDGI--Y 426
 QY 460 FIINQIEL-YLNGSNNLTLYKVSAGSLSNQNTFFQFPRKDCNVLVDPCSS--PNFNN 516
 Db 427 FGVTKVDFSQVDDQKNET-----STQYDSKRYNGYLGAQDSIDQLPPTTDEPLEKA 479
 QY 517 YSHILSHFSLFTYSYVIGLQILDTGVLGTHSSVDRYNAISDKIITMIPIAKGNLDT 576
 Db 480 YSHQLNVAECFL-----MQDRRGITIPFTWTHRSVDFNTIDAEKITQLPVVKAYALS 533
 QY 577 NSKVIEGPGHTGMLVYLQ-----SQRLIEITCETPNSTQSYFIRLRYATNGAGNLTNPIS 632
 Db 534 GASIEGPGTGGNLLFLKSSNSIAKFKVTLNSAALLQRYVRIRYAST-----TNLR 587
 QY 633 LTPGVIGIPQRLNNTFSGTNNYLNQY--GDFGY--FQFPSTVTLPLNRNIPFIENRAD 688
 Db 588 LFV-----QNSNDFLVIYINKTNWIDGLTYQTPDFATS-----NSNMGFSGDTND 634
 QY 689 -----VNSILIIDKIEFPI 704
 Db 635 FIIGAEFVSNEKIYIDKIEFIPV 658

RESULT 14

US-08-996-441B-38
 ; Sequence 38, Application US/08996441B
 ; Patent No. 6023013
 ; GENERAL INFORMATION:
 ; APPLICANT: English, Leigh H.
 ; APPLICANT: Brussock, Susan M.
 ; APPLICANT: Malvar, Thomas M.
 ; APPLICANT: Bryson, James W.
 ; APPLICANT: Kulesza, Caroline A.
 ; APPLICANT: Walters, Frederick S.
 ; APPLICANT: Slatin, Stephen L.
 ; APPLICANT: Von Tersch, Michael A.
 ; APPLICANT: Romano, Charles
 ; TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS
 ; NUMBER OF SEQUENCES: 113
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Arnold, White & Durkee
 ; STREET: P.O. Box 4433
 ; CITY: Houston
 ; STATE: Texas
 ; COUNTRY: USA
 ; ZIP: 77210
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/996,441B
 ; FILING DATE: 18-DEC-1997
 ; CLASSIFICATION: 800
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kitchell, Barbara S.
 ; REGISTRATION NUMBER: 33,928
 ; REFERENCE/DOCKET NUMBER: MECO:151
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 512/418-3000
 ; TELEFAX: 512/474-7577

; INFORMATION FOR SEQ ID NO: 38:

; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 652 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-996-441B-38

Query Match

Best Local Similarity 17.1%; Score 663; DB 2; Length 652;
 Best Local Similarity 28.1%; Pred. No. 8.3e-53;
 Matches 208; Conservative 114; Mismatches 292; Indels 126; Gaps 27;

QY 1 MNQNDNNEYIIDSHTSPYFENR--NSNDSVPYTNNEPQPLQNTNYKELNMCQ-GNT 57
 Db 1 MNPNN-RSEHD-----TIKVTNSELOTHNHQVPLADNPNSTLEELNYKEFLWEDSST 54
 QY 58 QYGDNFETPASADTTA-AVSAGTIVSGTLLAGIGGLTSISGPIIGIICAIISFGTLITVF 116
 Db 55 EVLDN-----STVKDAVGIGISVVGQILGVVG-----VPPAGALTSFYQSFLNTI 99
 QY 117 WPAGEQDKTVMTQFIKMGIEFVDTPLTBSIKOLKLTLEGFRQILQSYNTALDDWRKLR 176
 Db 100 WPS---DADPWKAFMAQVEVLIDKKIEEVAKSKALAEQLQVNFEDYVNALNSWK--- 153
 QY 177 LQAPGLPSSALQQAALTILKIRFENVHNDPIREIFQFQLETYKTLILLPIYAQAANFHLN 236
 Db 154 -----TPLSLRSKRSQDRIRELFSQAESHFRNMPFSAVSKFEVLFLPYAQAANTHLL 208
 QY 237 LQQAELADEWADLHPQIEBNAGTSDDYKLLKENIPKSYNYCANTYRTGLKNRDEP 296
 Db 209 LKDAQVGEEMGYS-----SEDAEFYHQQLKT-QQYTDHCNVNNGVNLGRGST 259
 QY 297 NMKWSIFNDYRYRYMTITVLDTISQFSLYDIKEVRDSIGGIEVKGIGKVELTRIIYITEINF 356
 Db 260 YDAWKFNFRREMTLTVLDLVLVLPFFVDRLYS-----KGKVELTRDIFTDPIFS 311
 QY 357 DRLPQLRVQPNLATMEYNLTASFLFSLFQFIFVT-----ENTNF--GNRLVGI 405
 Db 312 LRTP-LAYGPTFLSIENSIRKP--HLFDYLGIEFHTRLQPGYFGKDSFNYWSGN---V 365
 QY 406 SNRDAPTYNTITETLYGERTGSPPTKTIRPPES---YKVSIVTDROQSP-----VSP 455
 Db 366 ETRPSIGSKTITSPFYGDKSTEPVQKL--SPDGQKVYRTIANTDVAAMPNGKVYLGVTK 423
 QY 456 IQPHFIINQIELYLNNGSNNLTLYKVSAGSLSNQNTFFQFPRKDCNVLVDPCSSPNFN 515
 Db 424 VD-----FSQYDDQKNETSTQYDSKRNNGHVSQAQSIDQLPPTTD-----EPLEK 470
 QY 516 NYSHILSHFSLFTYSYVIGLQILDTGVLGTHSSVDRYNAISDKIITMIPIAKGNLND 575
 Db 471 AYSHQLNVAECFL-----MQDRRGITIPFTWTHRSVDFNTIDAEKITQLPVVKAYALS 524
 QY 576 TNSKVIEGPGHTGMLVYLQ-----SQRLIEITCETPNSTQSYFIRLRYATNGAGNLTNP 631
 Db 525 SGASIEGPGTGGNLLFLKSSNSIAKFKVTLNSAALLQRYVRIRYAST-----TNL 578
 QY 632 SLTIPGVIGIPQRLNNTFSGTNNY--NLOVGDGFGYFQFPSTVTLPLNRNIPFI 684
 Db 579 RLTV-----QNSNDFLVIYINKTNWIDGLTYQTPDLATNSNMGFSGDKNELIIG 630
 QY 685 NRADVNSILIIDKIEFPI 704
 Db 631 AESFVSNEKIYIDKIEFIPV 650

RESULT 15

US-08-993-722A-38
 ; Sequence 38, Application US/08993722A
 ; Patent No. 6060594
 ; GENERAL INFORMATION:
 ; APPLICANT: English, Leigh H.
 ; APPLICANT: Brussock, Susan M.
 ; APPLICANT: Malvar, Thomas M.

APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Tersch, Michael A.
APPLICANT: Romano, Charles
TITLE OF INVENTION: NUCLEIC ACID SEGMENTS ENCODING MODIFIED
COLLEOPTERAN-TOXIC CRYSTAL PROTEINS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESS: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,722A
FILING DATE: 18-DEC-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO.149
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3106
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 652 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-993-722A-38

Query Match 17.1%; Score 663; DB 2; Length 652;
Best Local Similarity 28.1%; Pred. No. 8.3e-53;
Matches 208; Conservative 114; Mismatches 292; Indels 126; Gaps 27;
QY 1 MNQNDNNEYIIDSHTSPYFNNR--NSNDSRYPTNNPNQPLQNTNYKEWLNMQ-GNT 57
DB 1 MNPNN-RSEHD-----TIKVTNSELQTNHQYPLADNPNSTLEELNYKEFLMTDSST 54
QY 58 QYGDNPETASADTIA-AVSAGTIVSGTLIAGTIGTISGPIGIGAIISFGTLITVF 116
DB 55 EVLDN-----STVKDAVGTGISVVGQILGVVG-----VPFAGALTSFYQSFLNTI 99
QY 117 WPAGEQDKTWTQFIKMGIFVDVTPLETISKQLQTLQLEGFRILQSYNTALDDWRKLR 176
DB 100 WPS---DADPWKAFMAQVEVLIDKKIEYAKSALAEQLQNNFEDYVNALSWKK--- 153
QY 177 LQAPGLPPSSALQQAALTUKIRENVNDPIREIPGQLETYKTLTLLPIYAAQANFHLNL 236
DB 154 -----TPLSLRSGRSQDRLELFSQAESHFRNSMPSFAVSKFEVLFPTTAAQANTHLL 208
QY 237 LQQAELADEWNADIHPSQIEPNAGTSDDYKLLKENIPKSYNCANTYRTGLKNLRDEP 296
DB 209 LKDAQVFGGEWGS-----SEDAEFVHRQLKT-QQYTDHCYNNVNVGLNGRGST 259
QY 297 NMKWSIFNDYRRYMTITVLDITQSPLDYDKRYRDSIGGIEVGKIKNELTREIYTTTEINF 356
DB 260 YDAWKFNFRFREMILTIVLPPFDIRLYS-----KGVKTELTRDIFTDPIFS 311
QY 357 DRLPOLRVQPNLATMEYNLTRASFKLPSLEQFIYV-----ENTNF--GNRLVGI 405
DB 312 LKTP-LAYGPTLUSIENSIRKP--HLFDLQGLIEFTRLPQPGYFGDKDYNWSGNY---V 365
QY 406 SNRDAPYTSNTITETLYGERTGTSPTTKTIRPFES---YKVSIVTDRQSPD-----VSP 455

Db 366 ETRPSIGSSKTIITSPPFYGDKSTPEVQKU--SFDQKVYRTTANTDVAAMPNGKVILGVTK 423
QY 456 IQPHFIINQIELYLNSSNNTLYKSAGGSLSNYQNTTFFQFPRKDCNLVIDPGGSPNPN 515
Db 424 VD---FSQYDDQKNETSTQYDSKRNGHVSAQDSIDQLPETTD-----EPLEK 470
QY 516 NYSHILSHFSLFTYVYVIGLQQLDGTGLGHTSHSVDRYNAISDKIITMIPAIKGNLJD 575
Db 471 AYSHQLNYAECFL-----MQDRRGTTIPFFTWTHRSVDFNTIDAEEKITQLPVPVKAYALS 524
QY 576 TNSKVIIEGPGHTGGLVYLQ-----SQGRLEITCETPNSTQSVFIRLRYATNGAGNTLPNI 631
Db 525 SGASIIIEGPGTGNLLFLKESNSIAKFKVLTLSAALLQRYRVIRYAST-----TNL 578
QY 632 SLATIPGVIGIPPPQRLNNTFSGTNYN-----NLQYGDGYFQFPSTVTTLPLNRNIPFIF 684
Db 579 RLFV-----QNSNNDFLVIYINKTMKDDDLTYQTFDLATTNSNMFGSGDKNELIIG 630
QY 685 NRADVNSIILIIDKIEFPI 704
Db 631 AESFVSNEKIVIDKIEFIPV 650

Search completed: December 4, 2005, 12:51:47
Job time : 52 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2005, 12:46:17 ; Search time 166 Seconds
(without alignments)
1850.027 Million cell updates/sec

Title: US-10-783-417-2
Perfect score: 3869
Sequence: 1 MNQNNNNEYEIDSHTSPV.....KLETIQTKINTFTNHTKYL 735

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.Main:*
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3869	100.0	735	5	US-10-783-417-2
2	3318	85.8	744	4	US-10-782-570-2
3	3104.5	80.2	694	4	US-10-782-570-4
4	1060.5	27.4	1180	4	US-10-782-141-12
5	1060.5	27.4	1180	4	US-10-782-096-14
6	1060.5	27.4	1180	4	US-10-782-570-10
7	1060.5	27.4	1180	5	US-10-783-417-8
8	1060.5	27.4	1180	5	US-10-781-979-14
9	892.5	23.1	675	4	US-10-782-141-17
10	892.5	23.1	675	4	US-10-782-096-18
11	892.5	23.1	675	4	US-10-782-570-14
12	892.5	23.1	675	5	US-10-783-417-12
13	892.5	23.1	675	5	US-10-781-979-19
14	892.5	23.1	675	5	US-10-926-819-16
15	849	21.9	1109	3	US-09-756-526A-4
16	849	21.9	1109	4	US-10-345-020-4
17	849	21.9	1109	4	US-10-342-821-4
18	758.5	19.6	1136	4	US-10-782-141-13
19	758.5	19.6	1136	5	US-10-781-979-15
20	758.5	19.6	1136	5	US-10-929-754-1
21	709	18.3	682	4	US-10-782-096-20
22	709	18.3	682	4	US-10-782-570-16
23	709	18.3	682	5	US-10-783-417-14
24	709	18.3	682	5	US-10-781-979-21
25	703	18.2	1210	4	US-10-032-717-4
26	703	18.2	1210	4	US-10-414-637-4
27	703	18.2	1210	4	US-10-606-320-4

28	703	18.2	1210	4	US-10-746-914-4	Sequence 4, Appli
29	689.5	17.8	1386	4	US-10-120-544A-6	Sequence 6, Appli
30	689.5	17.8	1386	6	US-11-091-654-6	Sequence 6, Appli
31	678.5	17.5	667	4	US-10-032-717-8	Sequence 8, Appli
32	678.5	17.5	667	4	US-10-414-637-8	Sequence 8, Appli
33	678.5	17.5	673	4	US-10-032-717-18	Sequence 18, Appli
34	678.5	17.5	673	4	US-10-414-637-18	Sequence 18, Appli
35	678.5	17.5	673	4	US-10-606-320-14	Sequence 14, Appli
36	678.5	17.5	673	4	US-10-746-914-14	Sequence 14, Appli
37	676	17.5	659	4	US-10-614-076-112	Sequence 112, App
38	676	17.5	659	4	US-10-782-096-12	Sequence 12, Appli
39	676	17.5	659	4	US-10-782-570-9	Sequence 9, Appli
40	676	17.5	659	5	US-10-783-417-7	Sequence 7, Appli
41	675	17.4	1206	4	US-10-032-717-2	Sequence 2, Appli
42	675	17.4	1206	4	US-10-414-637-2	Sequence 2, Appli
43	675	17.4	1206	4	US-10-606-320-2	Sequence 2, Appli
44	675	17.4	1206	4	US-10-746-914-2	Sequence 2, Appli
45	672.5	17.4	648	4	US-10-782-141-18	Sequence 18, Appli

ALIGNMENTS

RESULT 1
US-10-783-417-2
; Sequence 2, Application US/10783417
; Publication No. US20040216186A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargies, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; TITLE OF INVENTION: AXMI-006, A Delta-Endotoxin Gene and
; TITLE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/274146
; CURRENT APPLICATION NUMBER: US/10/783,417
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: 60/448,806
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-783-417-2

Query Match	100.0%	Score	3869	DB	5	Length	735
Best Local Similarity	100.0%	Pred. No.	1.6e-304				
Matches	735	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	MNQNNNNEYEIDSHTSPVFNNSNDSRYPYTNPNQPLQNTNYKSWLNMCCQNTQYG	60				
Db	1	MNQNNNNEYEIDSHTSPVFNNSNDSRYPYTNPNQPLQNTNYKSWLNMCCQNTQYG	60				
Qy	61	DNFETPASADTIAAVSAGTIVSGTLLAGIGLTSISGPIGIIIGAIISFGTLITVFWPAG	120				
Db	61	DNFETPASADTIAAVSAGTIVSGTLLAGIGLTSISGPIGIIIGAIISFGTLITVFWPAG	120				
Qy	121	EQDKTVMQTQIKMGEIFVDTPLTESIKQLKLTQLEGFQRILOSNTALDDWRKLRLQAP	180				
Db	121	EQDKTVMQTQIKMGEIFVDTPLTESIKQLKLTQLEGFQRILOSNTALDDWRKLRLQAP	180				
Qy	181	GLPSSALQQAALTILKIRFENVHNDFREIPGFOLETYKTLPIYAQAANFHLNLOOG	240				
Db	181	GLPSSALQQAALTILKIRFENVHNDFREIPGFOLETYKTLPIYAQAANFHLNLOOG	240				
Qy	241	AEADENWADHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKRLRDEPNMKW	300				
Db	241	AEADENWADHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKRLRDEPNMKW	300				
Qy	301	SIFNDYRRYMTITVLDITISQFSLYDIKRYRDSIGIEVKGIKNELTREIYTTTINFDRLP	360				

Db 301 SIENDYRRYMTITVLDITISQSLYDIKRYRDSIGGIEVKIGKNELTREIYTTINFDRLP 360
QY 361 QLRVQPNLATMEYNLTRASFLEQIFYYTENTNFGNRLVGIISNRDAPTYSNITET 420
Db 361 QLRVQPNLATMEYNLTRASFLEQIFYYTENTNFGNRLVGIISNRDAPTYSNITET 420
QY 421 LYGERTGSPPTKIRPFESYKVSIVTDROSPPVSPIQPHFIINQIELYLANGSNNTLKYS 480
Db 421 LYGERTGSPPTKIRPFESYKVSIVTDROSPPVSPIQPHFIINQIELYLANGSNNTLKYS 480
QY 481 AGSLSNYQNTFFQPPRKDCNVLDPGSPNPNYSHLSHFSFTYSYVIGLOQIL 540
Db 481 AGSLSNYQNTFFQPPRKDCNVLDPGSPNPNYSHLSHFSFTYSYVIGLOQIL 540
QY 541 DTGVLGTHSSVDRYNAISDKIITMIPAIGNNLDNNSKVIEGPGHTGGNVLVLSQGR 600
Db 541 DTGVLGTHSSVDRYNAISDKIITMIPAIGNNLDNNSKVIEGPGHTGGNVLVLSQGR 600
QY 601 EITCETPNSTQSYFIRLRYATNGAGNTLPNISLITIPGVIGIPQRLNNTFSGTYNNLQY 660
Db 601 EITCETPNSTQSYFIRLRYATNGAGNTLPNISLITIPGVIGIPQRLNNTFSGTYNNLQY 660
QY 661 GDFGVQFPSTVTLPLNRNIPFIENRADVSNLSIIIDKIEFIPITSSMHQNRKQKLETI 720
Db 661 GDFGVQFPSTVTLPLNRNIPFIENRADVSNLSIIIDKIEFIPITSSMHQNRKQKLETI 720
QY 721 IQTKINTFFTNHTK 735
Db 721 IQTKINTFFTNHTK 735

RESULT 2
US-10-782-570-2
; Sequence 2, Application US/10782570
; Publication No. US20040210965A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-007, A Delta-Endotoxin Gene and
; FILE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/274144
; CURRENT APPLICATION NUMBER: US/10/782,570
; CURRENT FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: 60/448,812
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 744
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-782-570-2

Query Match 85.8%; Score 3318; DB 4; Length 744;
Best Local Similarity 87.1%; Pred. No. 8.3e-260;
Matches 639; Conservative 31; Mismatches 60; Indels 4; Gaps 3;
QY 1 MNQNNNNEYEIIDSHTSPVFNPNRNSDKRYPYTNPNQPLQNTYKWLNMCCQNTQYG 60
Db 1 MNQNNNNEYEIIDSHTSPVFNPNRNSDKRYPYTNPNQPLQNTYKWLNMCCQNTQYG 59
QY 61 DNFETPASADTAAVSAAGTIVSGTLLAGIGLTSISGPIGIIIGALISFTLITVFPAG 120
Db 60 DNFETPASADTAAVSAAGTIVSGTLLAGIGLTSISGPIGIIIGALISFTLITVFPAG 119
QY 121 EQDKTWTQFIKMGEIFVDTPLTESIKQLKLTLEGFRQILQSYNTALDWRKLRLOAP 180
Db 120 EQDKTWTQFIKMGEIFVDTPLTESIKQLKLTLEGFRQILQSYNTALDWRKLRLOAP 179

QY 181 GLPSSALQQAALTTLKIRFENVHDFIREIPGFOLEYTKTLLLPYAAQANFHLNLLQOG 240
Db 180 GLPSSALQQAALTTLKIRFENVHDFIREIPGFOLEYTKTLLLPYAAQANFHLNLLQOG 239
QY 241 AELADENWADIHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKNLDRPNMKW 300
Db 240 AELADENWADIHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYREGNKLKRNPNRW 299
QY 301 SIENDYRRYMTITVLDITISQSLYDIKRYRDSIGGIEVKIGKNELTREIYTTINFDRLP 360
Db 300 SIENDYRRYMTITVLDITISQSLYDIKRYRDSIGGIEVKIGKNELTREIYTTINFDRLP 357
QY 361 QLRVQPNLATMEYNLTRASFLEQIFYYTENTNFGNRLVGIISNRDAPTYSNITET 420
Db 358 YLEIQPNLATMEYNLTRSGIRLFSFLDELIFYTKNETYGNRLVGIANRNRSTYATTGTEI 417
QY 421 LYGERTGSPPTKIRPFESYKVSIVTDROSPPVSPIQPHFIINQIELYLANGSNNTLKYS 479
Db 418 IYGERGPPPTKIRPFESYKVSIVTDROSPPVSPIQPHFIINQIELYLANGSNNTLKYS 477
QY 480 SAGSLSNYQNTFFQPPRKDCNVLDPGSPNPNYSHLSHFSFTYSYVIGLOQI 539
Db 478 SAGSLSNDKKTDFQPPVKCKCKPIINPNCLPSYNSYSHLSQSLFNYSYKIGLALNI 537
QY 540 LDTGVLGTHSSVDRYNAISDKIITMIPAIGNNLDNNSKVIEGPGHTGGNVLVLSQGR 599
Db 538 LYTALGTHSSVNRNNAISDKIITMIPAIGNNLDNNSKVIEGPGHTGGNVLVLSQGR 597
QY 600 LEITCETPNSTQSYFIRLRYATNGAGNTLPNISLITIPGVIGIPQRLNNTFSGTYNNLQ 659
Db 598 LEITCETPNSTQSYFIRLRYATNGAGNTLPNISLITIPGVIGIPQRLNNTFSGTYNNLQ 657
QY 660 YGDFGVQFPSTVTLPLNRNIPFIENRADVSNLSIIIDKIEFIPITSSMHQNRKQKLET 719
Db 658 YGDFGVQFPSTVTLPLNRNIPFIENRADVSNLSIIIDKIEFIPITSSMHQNRKQKLET 717
QY 720 IQTKINTFFTNHTK 733
Db 718 IQTKINTFFTNHTK 731

RESULT 3
US-10-782-570-4
; Sequence 4, Application US/10782570
; Publication No. US20040210965A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-007, A Delta-Endotoxin Gene and
; FILE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/274144
; CURRENT APPLICATION NUMBER: US/10/782,570
; CURRENT FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: 60/448,812
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 694
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-782-570-4

Query Match 80.2%; Score 3104.5; DB 4; Length 694;
Best Local Similarity 87.6%; Pred. No. 1.6e-242;
Matches 598; Conservative 30; Mismatches 52; Indels 3; Gaps 2;
QY 52 MCQNGTOYQGNFETPASADTAAVSAAGTIVSGTLLAGIGLTSISGPIGIIIGALISFT 111
Db 1 MCQNGTOYQGNFETPASADTAAVSAAGTIVSGTLLAGIGLTSISGPIGIIIGALISFT 60

Qy	112	LiTVfWPAGEODKtWtQFIKWGEIfVdTPLTESIKQLKtLQLEGPRLQSYNTALDDW	171
Db	61	LiTVfWPAGEODKtWtQFIKWGEIfVdTPLTESIKQLKtLQLEGPRLQSYNTALDDW	120
Qy	172	RKLKRLQAPGLPPSSALQQAALTLKIRFENVHDFIREIPGFOLEYTKtLLLPiVAQAAN	231
Db	121	RKLKRLQAPGLPPSSALQQAALTLKIRFENVHDFIREIPGFOLEYTKtLLLPiVAQAAN	180
Qy	232	FHLNLQQAELADEWNADIHPSQIEPNAGTSDDYYKLLKENIPKYSNYCANTYRTGLKN	291
Db	181	FHLNLQQAELADEWNADIHPSQIEPNAGTSDDYYKLLKENIPKYSNYCANTYRTGLKN	240
Qy	292	LRDEPNKWSIFNDYRRYWTtTVLDTISQFSLYDIKRYRDSGGIEVGKIKNELTREIYT	351
Db	241	LRNEPNRWSIFNDYRRYWTtTVLDTIAQFSYDIKRYKDSIG--RIGGIKTELTREIYT	298
Qy	352	TEINFDRLLPOLRVOPNLATMEYNLTIRASFCLPSFLEQFIYFYTENTNPNRLVGINRDP	411
Db	299	TEINFDRLLYLEIQNLAIMENLTIRSGULFSFLDELIFYTKNBTYGNRLVGINRNRKS	358
Qy	412	TYSNTITETLYGERTGSPtTKTIRPFESYKYSIVTDROSPPVSPi-QPHFIINQIELYLN	470
Db	359	TYATTGTIELYGERGTPTTKLTPPFESYKYSIVTDROVTSPSPPNiYFTINQIELYLN	418
Qy	471	GSSNNTLUKYSAGGSLSNYQNTTFFQFPRKDCNLVIDPGCGPNFNNYSHILSHFSLFTVS	530
Db	419	NSPSNKLTYSAGGSLNSDKKtTDfQFPVKDCCKPiINPCLPSYNSYSHILSQFSLEFVS	478
Qy	531	YVIGLQIQLDGTGLVGHWSHYDRYNAISDKIITWIPAKGNLDTNSKVIBGPCHTGN	590
Db	479	YKIGLALNilyTGALGTHSSVRNRNAISDKIITWIPAKGNSLDTNSKVIBGPCHTGN	538
Qy	591	LVYLQSGRLEITCETPNSTQSYfIRLRYATNGAGNTLPNLSTLPiGVIGIPQRLNNTF	650
Db	539	LVYLQSGRLEITCRTPNSTQSYfIRLRYATNGAGNTLPNLSTLPiGVIGIPQRLNNTF	598
Qy	651	SGTYNNNLQYDGFYFQPPSVTVTLPLNRNIPIFINRADVNSILiDKIEFIPtSSMHQ	710
Db	599	SGTYNNNLQYDGFYFQPPSVTVTLPLNENiPIFINRADVNSILiDKIEFIPtSSVRQ	658
Qy	711	NREKQKLETTQTKINTPFTNHTK	733
Db	659	NREKQKLETTQTKINTPFTNHTK	681

```

RESULT 4
US-10-782-141-12
; Sequence 12, Application US/10782141
; Publication No. US20040197917A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMT-014, A Delta-Endotoxin Gene and
; TITLE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/274143
; CURRENT APPLICATION NUMBER: US/10782,141
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: 60/448,632
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1180
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-782-141-12

```

Matches 271; Conservative 118; Mismatches 278; Indels 105; Gaps 24

Qy 1 MNQNNDNNEYIIDSHTSPYFPPNRNSDSDRYPTNNPNQLONTNYKEWLNMCQGNITQYG 60
Db :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||:::
1 MNPYNKNEYETLNASQKL--NILSNNTYRPIENSPKQLQSTNYKDNLNMCOQNOQYG 58
Qy 61 DNPEFPASADTIAAASGACTIVSGTLLAGIGGLTSISGPIGIGAIISIIFGTLLTVFWPAP 120
Db :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||:::
59 GDPEFTIDS---GELSAYTTIWGTVLTGFGETT-----PLGL---ALIGFGTLIPVLPFA 108
Qy 121 EODKTVMVTQFIKMGEIIPVDTPLETISKQLKQLTGLEGFRQILQSNTALTDDWRKLRKLQAP 180
Db :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||:::
109 DQSWT-WSDFITQTKNIIKKBIASTYSNANKILNRSFNVISHNHLKTWE-----NNP 162
Qy 181 GLPSSALOQAALTCLKIRPENVNDFIREIP--GFQLEYTKTLLPPIYAQAANPHMLLQ 238
Db :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||:::
163 NPQNTQDVRTQIQLVHVHFQNVIPELVNSCPPNPSCDYNNILVLSYAQAANLHLTVLN 222
Qy 239 QCAELADEWNADIHPQSIENAGTSDYYKLLKENPKYSNYCANYVRTGLKMLDEP-- 296
Db :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||:::
223 QAVKFEAYLKNNRFQDYLEP-LPTAIDYYPVLTKAIEDTYNCVTTYKKGSLNIKTTPDS 281
Qy 297 ---NMKWSIFNDVRRYMTITVLDTTSQSLSYDIKRYRDSIGGIEVKVGIKNELTRELYTT 352
Db :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||:::
282 NLDCGINWNTYRYKMTTAVLDLVALPFNYDVGYK-----PIGVQSELTRELIYQV 333
Qy 353 EIFNDRLPQLRVOPNLATMEYNLTRASFKLFSLEQIFY----TENTNFGR----- 401
Db :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||:::
334 -LNFESPYKY--DFQYQEDSLTRRP-HLFTWLDSLNFYEKAQTFPNEFTSHYMFHY 389
Qy 402 -LVGISNRDAPTYSNTITETLYXERTGSPPTKTI RPPESYKVSIVTDROSPVPSPQH 460
Db :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||:::
390 TLDNISOKSVFNGNHNVTDKL--KSLGLATNIYI-----F 422
Qy 461 IINQIEL---YLANGSSNNT-----LKYSAGSGLSNYO-NTTFQFPKKDC 502
Db :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||:::
423 LLNVI SLDNKYLDNYNINSKOMDFPINGTRLLEKELTAGSGQITYDNKNIFGLPILKRR 482
Qy 503 NLVDPGCSFNFNNYSHILSFSLFYVYVIGLQLDLTGVLGWTHSSHVDNRNAISDKI 562
Db :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||:::
483 ENQGNPLFTPDYNYSHILSFIKSLSPATYKTQVY-----TFAWTHSSVDPKNTIYTHL 537
Qy 563 ITMIPAIGNNLDTNKSVIBGGHGHTGNLVYLOSQGRLEITCTBPSNSTOSYFIRLRYATN 622
Db :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||:::
538 TTQIPAVKANSLGTASKVQVPGHGTGDLL--DFKHFKITCQHSNFQQSYFIRIRYASN 595
Qy 623 GAGNTLPNLSLITPGVIGIPBPQLNNTFSGTNNYNLOYGDPGYFQPSPSTVTLPLNRNIPF 682
Db :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||:::
596 GSANTRAVINLSIPGVAEL-GMALNPTFSGDTYTNLKXDFQYLEFSNEVKFAPNQNISL 654
Qy 683 IFNRADV-NSNIIIDKIEFIPITSSMHONREKOQLETTIOTKINTFTFNHTK 733
Db :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||::: :|||:::
655 VFNRSDVYNTVTLIDKIEFLPTRISIREDREREKOQLETVQOIIINTFYANPIK 706

RESULT 5
US-10-782-096-14
; Sequence 14, Application US/10782096
; Publication No. US20040210964A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargies, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-009, A Delta-Endotoxin Gene and
; FILE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/274148
; CURRENT APPLICATION NUMBER: US/10/782,096
; CURRENT FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: 60/448,633
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 23

```

RESULT 5
US-10-782-096-14
; Sequence 14, Application US/10782096
; Publication No. US20040210964A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duzel, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-009, A Delta-Endotoxin Gene and
; TITLE OF INVENTION: Methods for Its Use
; FILE REFERENCES: 045600/274148
; CURRENT APPLICATION NUMBER: US/10782,096
; CURRENT FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: 60/448,633
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 23

```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 1180
; TYPE: PRF
; ORGANISM: Bacillus thuringiensis
US-10-782-096-14

Query Match      27.4%; Score 1060.5; DB 4; Length 1180;
Best Local Similarity 35.1%; Pred. No. 2.2e-76;
Matches 271; Conservative 118; Mismatches 278; Indels 105; Gaps 24;

QY 1 MNQNDNNEYEIIIDSHTSPPYFNRNSNDSPYPTNNPQPLQNTNYKWLNMCOGNTQYG 60
Db 1 MPYQNKNEYETLNASQKLL--NISNNYTRYPIENSPKQLLOSTNYKDWLNMCOQNOQYG 58
QY 61 DNFEFPASADTTAAVSAGTIVSGTLLAGIGLTSISGPIGIIIGAIISFGTLITVFWPAG 120
Db 59 GDFETPIDS--GELSAITVVGTVLTGFGFT---PLGL---ALIGFTLIPVLFPQAQ 108
QY 121 EQDKTWTQFIKMGEIFVDTPLTESIKQLKLTQLEGRQILOSYNTALDWRKLRLOAP 180
Db 109 DOSNT--WSDFITQTKNIIKKEASTYISNANKILNRSFNVISTYHNHLKTWE-----NNP 162
QY 181 GLPPSSALQQAALTALKIRPENVHNDPIREIP--GFQLETYKTLILLPIYAAQANFHLNLLQ 238
Db 163 NPQNTQDVRTQIQLVHYHFQNVIPELVNSCPPNPSCDYINILVLSYAAQANLHLTVLN 222
QY 239 QCAELADEWNADIHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKNLRDEP-- 296
Db 223 QAVKEAYLKNRQPDYLEP-LPTAIDYYPVLTKAIEDYNTYCVTVYKKGNLIKTTPDS 281
QY 297 ----NMKWSIFNDYRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKIGKNELTREIYTT 352
Db 282 NLDGNNMNTYNTYRTKMTTAVLDLVALFPNDVGKY-----PIGVQSELTREIYQV 333
QY 353 EINFRLPOLRVQPNLATWEYNLTRASFLEQFIYF-----TENTNFGNR----- 401
Db 334 -LNFESPYYKY--DFQYQEDSLTRRP-HLFTWLDLSNFEYKAKQTTPNFFTSYNNMFHY 389
QY 402 -LVGISNRDAPYTSNTITETLYGERTGSPPTKTRIRPFESYKVSIVTDROQSPVPIQPHF 460
Db 390 TLDNISQKSSVFGNHNVTDKL--KSLGLATNIYI-----F 422
QY 461 IINQIEL---YLANGSSNNT-----LKYSAGGSLSNYQ-NTTFFQPRKDC 502
Db 423 LLNVISLDNKYLNDYNNISKMPFFITNGTRLEKEKELTAGSQITYDVNKNIFGLPILKRR 482
QY 503 NLVIDPGCSPNFNNYSHILSHFSLFYSYVIGLQQLDGTGLVGTWTHSSVDYRNAISDKI 562
Db 483 ENQGNPTLPPTYDNYSHILSFISLIPATYKTQVY-----TFAWTHSSVDPKNTIYTHL 537
QY 563 ITWIPAIGKNNLDTSKVIETGPGHTGGLNVLVLOSQRLEITCETPNSTQSYFIRLYATN 622
Db 538 TTQIPAVKANSLGTASKVVGPGHTGGDLI--DFKDHFKITCOHSNFOQSYFIRIYASN 595
QY 623 GAGNTLPNLSLTPGVIGIPQORLNTFSGTNNYNNLOYGDGFGYFQFPSTVTLPLNRNIPP 682
Db 596 GSANTRAVINLSIPGVAEL--GMALNPTFSGTDYTNLKYKDFQYLEFNEVKFAPNQISL 654
QY 683 IFNRADV--SNSIILIDKIEFIPITSSMHQNRKQKLETTQTKINTFTNHTK 733
Db 655 VFNRSDVYNTTTLVDIKIEFLPITRSIREDEKQKLETVQOIINTFYANPIK 706
```

RESULT 6

```
US-10-782-570-10
; Sequence 10, Application US/10782570
; Publication No. US20040210965A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
```

```
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-007, A Delta-Endotoxin Gene and
; FILE REFERENCE: 045600/274144
; CURRENT APPLICATION NUMBER: US/10/782,570
; CURRENT FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: 60/448,812
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1180
; TYPE: PRF
; ORGANISM: Bacillus thuringiensis
US-10-782-570-10
```

```
Query Match      27.4%; Score 1060.5; DB 4; Length 1180;
Best Local Similarity 35.1%; Pred. No. 2.2e-76;
Matches 271; Conservative 118; Mismatches 278; Indels 105; Gaps 24;

QY 1 MNQNDNNEYEIIIDSHTSPPYFNRNSNDSPYPTNNPQPLQNTNYKWLNMCOGNTQYG 60
Db 1 MPYQNKNEYETLNASQKLL--NISNNYTRYPIENSPKQLLOSTNYKDWLNMCOQNOQYG 58
QY 61 DNFEFPASADTTAAVSAGTIVSGTLLAGIGLTSISGPIGIIIGAIISFGTLITVFWPAG 120
Db 59 GDFETPIDS--GELSAITVVGTVLTGFGFT---PLGL---ALIGFTLIPVLFPQAQ 108
QY 121 EQDKTWTQFIKMGEIFVDTPLTESIKQLKLTQLEGRQILOSYNTALDWRKLRLOAP 180
Db 109 DOSNT--WSDFITQTKNIIKKEASTYISNANKILNRSFNVISTYHNHLKTWE-----NNP 162
QY 181 GLPPSSALQQAALTALKIRPENVHNDPIREIP--GFQLETYKTLILLPIYAAQANFHLNLLQ 238
Db 163 NPQNTQDVRTQIQLVHYHFQNVIPELVNSCPPNPSCDYINILVLSYAAQANLHLTVLN 222
QY 239 QCAELADEWNADIHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKNLRDEP-- 296
Db 223 QAVKEAYLKNRQPDYLEP-LPTAIDYYPVLTKAIEDYNTYCVTVYKKGNLIKTTPDS 281
QY 297 ----NMKWSIFNDYRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKIGKNELTREIYTT 352
Db 282 NLDGNNMNTYNTYRTKMTTAVLDLVALFPNDVGKY-----PIGVQSELTREIYQV 333
QY 353 EINFRLPOLRVQPNLATWEYNLTRASFLEQFIYF-----TENTNFGNR----- 401
Db 334 -LNFESPYYKY--DFQYQEDSLTRRP-HLFTWLDLSNFEYKAKQTTPNFFTSYNNMFHY 389
QY 402 -LVGISNRDAPYTSNTITETLYGERTGSPPTKTRIRPFESYKVSIVTDROQSPVPIQPHF 460
Db 390 TLDNISQKSSVFGNHNVTDKL--KSLGLATNIYI-----F 422
QY 461 IINQIEL---YLANGSSNNT-----LKYSAGGSLSNYQ-NTTFFQPRKDC 502
Db 423 LLNVISLDNKYLNDYNNISKMPFFITNGTRLEKEKELTAGSQITYDVNKNIFGLPILKRR 482
QY 503 NLVIDPGCSPNFNNYSHILSHFSLFYSYVIGLQQLDGTGLVGTWTHSSVDYRNAISDKI 562
Db 483 ENQGNPTLPPTYDNYSHILSFISLIPATYKTQVY-----TFAWTHSSVDPKNTIYTHL 537
QY 563 ITWIPAIGKNNLDTSKVIETGPGHTGGLNVLVLOSQRLEITCETPNSTQSYFIRLYATN 622
Db 538 TTQIPAVKANSLGTASKVVGPGHTGGDLI--DFKDHFKITCOHSNFOQSYFIRIYASN 595
QY 623 GAGNTLPNLSLTPGVIGIPQORLNTFSGTNNYNNLOYGDGFGYFQFPSTVTLPLNRNIPP 682
Db 596 GSANTRAVINLSIPGVAEL--GMALNPTFSGTDYTNLKYKDFQYLEFNEVKFAPNQISL 654
QY 683 IFNRADV--SNSIILIDKIEFIPITSSMHQNRKQKLETTQTKINTFTNHTK 733
Db 655 VFNRSDVYNTTTLVDIKIEFLPITRSIREDEKQKLETVQOIINTFYANPIK 706
```

```
RESULT 7
US-10-783-417-8
; Sequence 8, Application US/10783417
; Publication No. US20040216186A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargies, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-006, A Delta-Endotoxin Gene and
; TITLE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/274146
; CURRENT APPLICATION NUMBER: US/10783.417
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: 60/448,806
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1180
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-783-417-8

Query Match      27.4%; Score 1060.5; DB 5; Length 1180;
Best Local Similarity 35.1%; Pred. No. 2.2e-76;
Matches 271; Conservative 118; Mismatches 278; Indels 105; Gaps 24;

Qy      1 MNQNDNNEYIIDSHTSPYFPNRSNDSRYPTNNPQLONTNYKEWLNKCOQNTQYG 60
Db      1 MNPQNKNYEYELNASOKKL--NISNNYTRYPIENSQKLLQSTNYKDWLNKCOQNTQYG 58

Qy      61 DNFTFASADTIAAVSAGTIVSGTLLAGIGLTSISGPIGIIIGAIISFGTLIVFWPAG 120
Db      59 GDFETFIDS---GELSAYTIIVGTVLTGFGFTT---PLGL---ALIGFGTLIPVLPFAQ 108

Qy      121 EQDKTVMTQFIKMGEIFVDPTLTESI KQLKLTLEGFRQILQSYNTALDWRKLRLOAP 180
Db      109 DQSNLT-WSDFITQTKNIKKIEIASTYISNANKILNRSFNVISTYHNHLKTWE-----NNP 162

Qy      181 GLPPSSALQQAALTLKIRFENVHNDIFREIP--GFQLETYKTLTLLPIYAAANFHLNLLQ 238
Db      163 NPQNTQDVRTQIQLVHVFQNVIPELVNSCPNPSDCDYNNILVSSYAAANLHLTVLN 222

Qy      239 QGAELADEWNADTHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKNLRDEP-- 296
Db      223 QAVKFEAYLKNRQFDYLEP-LPTAIDYYPVLTKAIEDYNYCVTYTKKGLNLIKTTTDS 281

Qy      297 ----NMKWSIFNDYRRYNTITVLDTISQFSLYDIKRYRDSIGGIEVGKIKNELTREIYTT 352
Db      282 NLDGNINWNTYRTKMTTAVLDLVALFPNDVGKY-----PIGVQSELTREIYQV 333

Qy      353 EINFDRPLQRLVQPNLATMEYNLTRASPKLSFLEQIFY-----TENTNFGNR----- 401
Db      402 -LNFEEPSYKY--DFQYQEDSLTRRP-HLFTWLDLSLNFYEKAQTTNNFTSHYMFHY 389

Qy      402 -LVGIGNRDAPTYSNITITETLYGERTGSPPTKTRIRPFESYKVSIVTDRQSPVPSPIQPHF 460
Db      390 TLDNISQKSSVFGNHNVTDKL--KSLGLATNIYI-----F 422

Qy      461 IINQIEL---YLNGSSNNT-----LKYSAGGSLSNYQ--NTTFFQPPRKKDC 502
Db      423 LLNVISLDNKYLDYNNISKMDFFITNGTRLLLEKELTAGSGQIYDYNKIFGLPIKRR 482

Qy      503 NLVIDGCSNPNFNYSILSHFSIFTYSYVIGLQILDGVLGWTHSSVDYRNAISDKI 562
Db      483 ENQGNPTLFFTYDNYSHLSFIKSLSPATYKTVQV-----TPAWTHSSVDPKNTIYTHL 537

Qy      563 ITMPIAKGNLDTNSKVISGPGHTGNLVYLOSQGRLEITCTETPNSTQSYFPLRVATN 622
Db      538 TTQIPAVKANSLGATASKVVGPGHTGDLI--DPKDHFKITCQHSNFPQSQSYFIRIVASN 595
```

```
Qy      623 GAGNTLPNLSLTIRGVIGIPQORLNNSTGTYNNNLOYGDFGYFQFPSPSTVTLPLNRP 682
Db      596 GSANTRAVINUSIPGVAEL-GMALNPTSGDYTNLKYDKDFQYLFBSNEVKAPNQNISL 654

Qy      683 IFNRADV-SNSILIIDKIEFIPITSSMHQNKREKQKLETIQTKINTFFTHHTK 733
Db      655 VFNRSVDVYNTTVLIDKIEFLPITRSIREDEKQKLETVOQIINTFYANPIK 706

RESULT 8
US-10-781-979-14
; Sequence 14, Application US/10781979
; Publication No. US20040250311A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargies, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-008, A Delta-Endotoxin Gene and
; TITLE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/274147
; CURRENT APPLICATION NUMBER: US/10781.979
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: 60/448,797
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 1180
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-781-979-14

Query Match      27.4%; Score 1060.5; DB 5; Length 1180;
Best Local Similarity 35.1%; Pred. No. 2.2e-76;
Matches 271; Conservative 118; Mismatches 278; Indels 105; Gaps 24;

Qy      1 MNQNDNNEYIIDSHTSPYFPNRSNDSRYPTNNPQLONTNYKEWLNKCOQNTQYG 60
Db      1 MNPQNKNYEYELNASOKKL--NISNNYTRYPIENSQKLLQSTNYKDWLNKCOQNTQYG 58

Qy      61 DNFTFASADTIAAVSAGTIVSGTLLAGIGLTSISGPIGIIIGAIISFGTLIVFWPAG 120
Db      59 GDFETFIDS---GELSAYTIIVGTVLTGFGFTT---PLGL---ALIGFGTLIPVLPFAQ 108

Qy      121 EQDKTVMTQFIKMGEIFVDPTLTESI KQLKLTLEGFRQILQSYNTALDWRKLRLOAP 180
Db      109 DQSNLT-WSDFITQTKNIKKIEIASTYISNANKILNRSFNVISTYHNHLKTWE-----NNP 162

Qy      181 GLPPSSALQQAALTLKIRFENVHNDIFREIP--GFQLETYKTLTLLPIYAAANFHLNLLQ 238
Db      163 NPQNTQDVRTQIQLVHVFQNVIPELVNSCPNPSDCDYNNILVSSYAAANLHLTVLN 222

Qy      239 QGAELADEWNADTHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKNLRDEP-- 296
Db      223 QAVKFEAYLKNRQFDYLEP-LPTAIDYYPVLTKAIEDYNYCVTYTKKGLNLIKTTTDS 281

Qy      297 ----NMKWSIFNDYRRYNTITVLDTISQFSLYDIKRYRDSIGGIEVGKIKNELTREIYTT 352
Db      282 NLDGNINWNTYRTKMTTAVLDLVALFPNDVGKY-----PIGVQSELTREIYQV 333

Qy      353 EINFDRPLQRLVQPNLATMEYNLTRASPKLSFLEQIFY-----TENTNFGNR----- 401
Db      334 -LNFEEPSYKY--DFQYQEDSLTRRP-HLFTWLDLSLNFYEKAQTTNNFTSHYMFHY 389

Qy      402 -LVGIGNRDAPTYSNITITETLYGERTGSPPTKTRIRPFESYKVSIVTDRQSPVPSPIQPHF 460
Db      390 TLDNISQKSSVFGNHNVTDKL--KSLGLATNIYI-----F 422

Qy      461 IINQIEL---YLNGSSNNT-----LKYSAGGSLSNYQ--NTTFFQPPRKKDC 502
Db      423 LLNVISLDNKYLDYNNISKMDFFITNGTRLLLEKELTAGSGQIYDYNKIFGLPIKRR 482
```


Db 212 NLLKHAATYNIW---LQNGINPSTFNSNYQGLKRIQYDYTCIQYNAGLTWIR 268
Qy 294 DEPNMKWSIFNDYRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKGKINELTREIYYTE 353
Db 269 TTNATNMNTYRLEMTLTVLDLIAIPPNYDPEKY-----PIGVKSELIREVY-TN 319
Qy 354 INFRLPQLRVQPNLATMEYNLTIRASPKLFSFLQFIYFTENTN-----FGNRL 402
Db 320 VNSDTF-----RTITELNGLTR-NPTLFTWINQGRFYTRNSRDILDPYDFISFTGNQM 372
Qy 403 VGISNRDAPYNSNITITLYGERTG---SPTTKTIRPPESY---KVSIVTDROSPPVSP 455
Db 373 -----AFTHNDNRNIWGAVHGNIIISQDTSKVFPYRNKPIDKVEIVRREYSDIIL- 424
Qy 456 IQPHFIINQIELYNGSNNTLKYSAGSL-SNYQNTTFFQPPKKDCNVLDPGCSPNF 514
Db 425 -----YEMIFPNSSEVFYSSNSTIENNYKRTDSYMPKQTKN----- 464
Qy 515 NNYSHILSHFSLFTYSYVIGLQILDTGLVGLWTHSSVDYRNAISDKIITMIPAIGNNL 574
Db 465 EYGHITLSYIKTDNYIFSVVRERRV---AFSWHTSVDFTQNTIDLDNITQIHAKALKV 521
Qy 575 DTNSKVLGPGHGTGGLNVLVLOSQGRLEITCETPNSTQSYFIRLRYATGAGNTLPNISLT 634
Db 522 SSDSKIYKVGPGHGTGGLVILKDSMDFVRFP-LKNVSRQYQVRIYATNA-----PKTTFV 575
Qy 635 IPGVIGIPQRLNNTFFSGTNNY--NLQYDGFYGFQFPSTV-----TLPLNRNIPFI 683
Db 576 LTGIDTISVE-LPSTTSRQNPATDLTYADFGYVTFPRTVPNKTFEGEDTLMT-----L 629
Qy 684 FNRAVNSNLIIDKIEFIPITSSMHQNRKQKLEITQTKINTFTFN 730
Db 630 YGTPNHSYNI-YIDKIEFIPITQSVDLYTEKQNIETQKIVNDLFVN 675

RESULT 11

US-10-782-570-14
; Sequence 14, Application US/10782570
; Publication No. US20040210965A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-007, A Delta-Endotoxin Gene and
; FILE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/274144
; CURRENT APPLICATION NUMBER: US/10782,570
; PRIOR FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: 60/448,812
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-782-570-14

Query Match 23.1%; Score 892.5; DB 4; Length 675;
Best Local Similarity 31.9%; Pred. No. 4e-63;
Matches 245; Conservative 126; Mismatches 267; Indels 129; Gaps 30;

Qy 1 MNQNDNNEVEIIDSHTSPYFPNRSNDSRYPTNNQPLQNTNYKWLNMCCQNTQYG 60
Db 1 MNPYNKNEVEIFNAPNGF--SKSNYSRYPLANKNPQKNTNYKDWLNVCDNQOYG 58
Qy 61 DNFTFASADTIAVASGTVISGILLGGLTSGISGPIGIIIGAILIISFGILITVFPAG 120
Db 59 NNAGNFASSETIVGSAGIIVVGTMGLG-----AFAAP--VLAAGIISFGTLPIFW-QG 109

Qy 121 BODKTVTQFIKMGEIFVDTPLTE---SIKQLKQLTLEGFRQILQSYNTALDDMRKLKRL 177
Db 110 SDPANWQDLNIG---GRPIQIEDKNILNVLTSIVTPIKNQLDKYQEFDFKWEPAF-- 163
Qy 178 QAPGLPSSALQQAALTILKIRFENVHN---DFIREIPGFOLEYTKTLLPIYAAQANFHL 234
Db 164 -----THANAKAVHDLFTTLEPIIDKDLMLKNNASVRIPT-----LPAYAQIATWHL 211
Qy 235 NLLQGAELAEWNAADIHPSQIEPNAGTSDDYK--LLKENIPKYSNYCANTYRTGLKNLR 293
Db 212 NLLKHAATYNIW---LQNGINPSTFNSNYQGLKRIQYDYTCIQYNAGLTWIR 268
Qy 294 DEPNMKWSIFNDYRYMTITVLDTISQFSLYDIKRYRDSIGGIEVKGKINELTREIYYTE 353
Db 269 TTNATNMNTYRLEMTLTVLDLIAIPPNYDPEKY-----PIGVKSELIREVY-TN 319
Qy 354 INFRLPQLRVQPNLATMEYNLTIRASPKLFSFLQFIYFTENTN-----FGNRL 402
Db 320 VNSDTF-----RTITELNGLTR-NPTLFTWINQGRFYTRNSRDILDPYDFISFTGNQM 372
Qy 403 VGISNRDAPYNSNITITLYGERTG---SPTTKTIRPPESY---KVSIVTDROSPPVSP 455
Db 373 -----AFTHNDNRNIWGAVHGNIIISQDTSKVFPYRNKPIDKVEIVRREYSDIIL- 424
Qy 456 IQPHFIINQIELYNGSNNTLKYSAGSL-SNYQNTTFFQPPKKDCNVLDPGCSPNF 514
Db 425 -----YEMIFPNSSEVFYSSNSTIENNYKRTDSYMPKQTKN----- 464
Qy 515 NNYSHILSHFSLFTYSYVIGLQILDTGLVGLWTHSSVDYRNAISDKIITMIPAIGNNL 574
Db 465 EYGHITLSYIKTDNYIFSVVRERRV---AFSWHTSVDFTQNTIDLDNITQIHAKALKV 521
Qy 575 DTNSKVLGPGHGTGGLNVLVLOSQGRLEITCETPNSTQSYFIRLRYATGAGNTLPNISLT 634
Db 522 SSDSKIYKVGPGHGTGGLVILKDSMDFVRFP-LKNVSRQYQVRIYATNA-----PKTTFV 575
Qy 635 IPGVIGIPQRLNNTFFSGTNNY--NLQYDGFYGFQFPSTV-----TLPLNRNIPFI 683
Db 576 LTGIDTISVE-LPSTTSRQNPATDLTYADFGYVTFPRTVPNKTFEGEDTLMT-----L 629
Qy 684 FNRAVNSNLIIDKIEFIPITSSMHQNRKQKLEITQTKINTFTFN 730
Db 630 YGTPNHSYNI-YIDKIEFIPITQSVDLYTEKQNIETQKIVNDLFVN 675

RESULT 12

US-10-783-417-12
; Sequence 12, Application US/10783417
; Publication No. US20040216186A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-006, A Delta-Endotoxin Gene and
; FILE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/274146
; CURRENT APPLICATION NUMBER: US/10783,417
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: 60/448,806
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-783-417-12

Query Match 23.1%; Score 892.5; DB 5; Length 675;
Best Local Similarity 31.9%; Pred. No. 4e-63;
Matches 245; Conservative 126; Mismatches 267; Indels 129; Gaps 30;

; TITLE OF INVENTION: AXMI-003, A Delta-Endotoxin Gene and
; FILE REFERENCE: Methods for Its Use
; CURRENT APPLICATION NUMBER: US/10/926,819
; CURRENT FILING DATE: 2004-08-26
; PRIOR APPLICATION NUMBER: 60/498,518
; PRIOR FILING DATE: 2003-08-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-926-819-16

Query Match 23.1%; Score 892.5; DB 5; Length 675;
Best Local Similarity 31.9%; Pred. No. 4e-63;
Matches 245; Conservative 126; Mismatches 267; Indels 129; Gaps 30;
QY 1 MNQNNNVEIIDSHTSPYFPNNSDSRYPTNNNOPLONTNYKEWLNMCQNTQYG 60
Db 1 MNPYQNKEYEINAPSGNF--SKSNYSRYPLANKNPQKNTNYKDWLNVCDNQOY 58
QY 61 DNPETASADTAAVASGTVSGTLLAGIGLTSISGPIGIIIGAIISFGTLITVFWPAG 120
Db 59 NNAGNFASSETIVGVSAGIIVGTMLG-----AFAAP--VLAAGIISFGTLLPIFW-QG 109
QY 121 EODKTVMQFIKGEIFVDTPLTE---SIKQLKLTLEGFRQILOSNTALDDWRKLR 177
Db 110 SDPANWQDLLNIG---GRPIQIDKNIINVLTSIVTPIKNQDKQEFDFKWEPA-- 163
QY 178 QAPGLPSSALQQAALTKTRFENVN---DFIREIFGQLETYKTLPIYAAQANFHL 234
Db 164 -----THAKAVHDLFTLEPIIDKOLDMLKNNASYRIPT-----LPAYAIATWHL 211
QY 235 NLLQOQGAELADENADHPQIEBNAGTSDDYK-LKKNIPKYSNYCANTYRTGLKNLR 293
Db 212 NLLKHAATYNIW---LONQGINPTNSNYYQYIKRIQIYDTCIOTYNAGLTMR 268
QY 294 DEPNKWSIFNDYRRYNTITVDTISOFSLYDIKRYRDSIGGIEVKGIKNELTREIYTT 353
Db 269 TWTNATWNTYRTLENTLVLDLIAFPNYDEPKY-----PIGVKSELIREVY-TN 319
QY 354 INFDRLPQURVQNLMATMEYNLTASPKLPSFLBQFIPTENTN-----FGNRL 402
Db 320 VNSDTE-----RTITELNGLTR-NPTLFTWINGQRFYTRNSRDILDPYDFISFTGNQM 372
QY 403 VGISNRDAPYSNTITETLYGERTG---SPTTKTIRPFESY---KYSIVTDROSPPVSP 455
Db 373 -----AFHTNDNRNIIWGAHVGNIIISQDTSKVFPFPRNKPDKVEIVRHRREYS 424
QY 456 IQPHFIINQIELYNGSSNNTLYKSAGGSL-SNYQNTTFQFPRKDCNLEVIDPGCSNF 514
Db 425 -----YEMIFFNSSESVFYSSNTIENNYKRTDSTWIPQKTWN----- 464
QY 515 NNYSHILSHFSLTYSYVIGLQILPTGLVGTWTHSSVDRYNAISDKIITMIPIAKGNL 574
Db 465 EEEGHTLSYKTONYIFSVVRERRV---AFSWHTTSVDQNTIDLONITQIHALKALKV 521
QY 575 DTNKSIVIEGHTGGMNLYVLOSQRLEITCETPNSTQSYFIRURYANGAGNTLPINLSLT 634
Db 522 SSDSKIVKPGHGTGGDLVILKDSMDFRVRF-LKNVSRQYQVRIYATNA-----PKTTFV 575
QY 635 IPGVIGIPPORLANTTSGTNYN--NLOYGDFYFQFPSTV-----TLPLNRNIPFI 683
Db 576 LTGIDITISVE-LPSTTSRQPNATDLYADFGYVTFPTVPNTKTEGEDTLLMT-----L 629
QY 684 ENRADVSNSTLIIDKTEFIPITSSMHQNRKOKLETIQTKNTPFTN 730
Db 630 YGTPNHSYNI-YIDKTEFIPITOSVLDYETBKQNIETKQKIVNOLFVN 675

RESULT 15

US-09-756-526A-4
; Sequence 4, Application US/09756526A
; Patent No. US20020038005A1
; GENERAL INFORMATION:
; APPLICANT: Jana, Wojciechowska
; APPLICANT: Evgeny, Lewitin
; APPLICANT: Ludmila, Revina
; APPLICANT: Igor, Zalunin
; APPLICANT: Galina, Chestukhina
; TITLE OF INVENTION: NOVEL DELTA-ENDOTOXINS AND NUCLEIC ACID SEQUENCES CODING THEREFOR,
; FILE REFERENCE: S-30913A
; CURRENT APPLICATION NUMBER: US/09/756,526A
; CURRENT FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: US 60/175,158
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 4
; LENGTH: 1109
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-09-756-526A-4

Query Match 21.9%; Score 849; DB 3; Length 1109;
Best Local Similarity 33.3%; Pred. No. 2.9e-59;
Matches 229; Conservative 106; Mismatches 248; Indels 104; Gaps 25;
QY 68 SADTIAAVSAGTIVSGTLLAGIGLTSISGPIGIIIGAIISFGTLITVFWPAGQD-KTV 126
Db 28 SSDIVAVVSAGIVVVGTTLT---APASFVNP---GVVLISFGTLAPVLPDPEEDPKKI 80
QY 127 WTQIKMGEIFVDTPLTESIKQLKLTLEGFRQILOSNTALDDWRKLRQAQGLPPSS 186
Db 81 WSPQWKHGEDLLNQTISTAVKEIALAHLNGFKDLVTIYERAFENDWR-----NPSA 131
QY 187 ALQQAALTKTRFENVNDFFIREIPGFOLEYTKTLPIYAAQANFHLNLLQGAELADE 246
Db 132 ---NTARLVSORFENAHFNFSVNNPQLQPTVDTLLSCYTEAANLHLNLLHQQGVQFADQ 188
QY 247 WNADIHPESQIBPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKNLRDPNNKWSIFNDY 306
Db 189 WNADQHPSPMLKSSGT---YYDELLVYIEKINYCTKYHKLHKLKSEKITWDAYNTY 245
QY 307 RRYMTITVDTISOFSLYDIKRYRDSIGGIEVKGIKNELTREIYTTINFDRLPQURVOP 366
Db 246 RREMTLIVDLVATFPFYDIRRF-----PRGVELELTREVYTSLDHLTRPP----- 291
QY 367 NLATMEYNLTASPKLPSFLBQFIPTENTWFGNPLVGISNRDAPYSNT---ITETLYG 423
Db 292 -----GLFTWLSDIELYTESVAGDYLSGI--RESKYITGNGQFTWKNIY 335
QY 424 ERTG-SPTTKTIRPFESYKYSIVTDROSPPVSPPIQPHFIINQIELYNGSSNNTLYKSAG 482
Db 336 NTNRLSKQLITLLPGE-FWTHLSINRPQTITAGINKLSLQIKIVFTTFKNDN--EYQKN 392
QY 483 GSLSNY---QMTTFQFPRKDCNLEVIDPGCSNPNFNYSILSHFSLFTSYVIGLQLOI 539
Db 393 FNVANNQEPQETT-----NYPNDYGGG-NSQKFKHNLSHFPLIHH-----KLEF 435
QY 540 LD-----TGVLGWTHSSVDRYNAISDKIITMIPIAKGNLDTNKSIVIEGHTGGMNLYVLO 595
Db 436 AEYPHSPFALGWTHTNSVNSQNLISSESVSQTPLVKAYEV-TNNSVIRGSGFTGGDLIELR 494
QY 596 SQGRLEITCETPNSTQSYFIRURYANGAGNTLPINLSLTIPGVIGIPPORLANTTSGTNY 655
Db 495 D--KCSIKCKA-SSLKKYAISLFYAANNAIAVSDVGDGSGAGVL-----LOFTFSRKN 545
QY 656 NN-----LOYGDFYFQFPSTVTLPLNRNIPFIENRAD--VNSNLIIDKIEFIPITSSM 708
Db 546 NNFTIQDLNPKDFQVHTLLVDIELPESEIHLKREDDYEGVILLIDKLEFKPIDENY 605
QY 709 HQNRKOKLETIQTKNTPFTNHTKTL 735

Db 606 ---TNEMLKAKKAVNVLFINATNAL 629

Search completed: December 4, 2005, 13:04:00
Job time : 169 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2005, 12:51:02 ; Search time 11 Seconds
(without alignments)

319.948 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869

Sequence: 1 MQQNNNEYIIDSHSPY.....KLETIQKINTFFNHTKYL 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 26661 seqs, 4788334 residues

Total number of hits satisfying chosen parameters: 26661

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New.*

- 1: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	703	18.2	1210	7	US-11-058-727-4
2	703	18.2	1210	7	US-11-058-727-4
3	689.5	17.8	1386	7	US-11-091-643-6
4	678.5	17.5	673	7	US-11-058-727-14
5	678.5	17.5	673	7	US-11-058-727-14
6	675	17.4	1206	7	US-11-058-727-2
7	675	17.4	1206	7	US-11-058-727-2
8	669	17.3	1316	7	US-11-091-643-4
9	657.5	17.0	675	7	US-11-058-727-74
10	657.5	17.0	675	7	US-11-058-727-80
11	657.5	17.0	675	7	US-11-058-727-74
12	657.5	17.0	675	7	US-11-058-727-80
13	657	17.0	674	7	US-11-058-727-82
14	657	17.0	674	7	US-11-058-727-82
15	656.5	17.0	675	7	US-11-058-727-42
16	656.5	17.0	675	7	US-11-058-727-42
17	656.5	17.0	675	7	US-11-058-727-42
18	656.5	17.0	675	7	US-11-058-727-42
19	656	17.0	674	7	US-11-058-727-50
20	656	17.0	674	7	US-11-058-727-50
21	654.5	16.9	673	7	US-11-058-727-70
22	654.5	16.9	673	7	US-11-058-727-70
23	653.5	16.9	673	7	US-11-058-727-34
24	653.5	16.9	673	7	US-11-058-727-68
25	653.5	16.9	673	7	US-11-058-389-34

26	653.5	16.9	673	7	US-11-108-389-68	Sequence 68, Appl
27	653	16.9	674	7	US-11-058-727-76	Sequence 76, Appl
28	653	16.9	674	7	US-11-108-389-76	Sequence 76, Appl
29	652.5	16.9	673	7	US-11-058-727-22	Sequence 22, Appl
30	652.5	16.9	673	7	US-11-058-727-64	Sequence 64, Appl
31	652.5	16.9	673	7	US-11-058-727-66	Sequence 66, Appl
32	652.5	16.9	673	7	US-11-108-389-22	Sequence 22, Appl
33	652.5	16.9	673	7	US-11-108-389-64	Sequence 64, Appl
34	652.5	16.9	673	7	US-11-108-389-66	Sequence 66, Appl
35	652.5	16.9	675	7	US-11-058-727-78	Sequence 78, Appl
36	652.5	16.9	675	7	US-11-108-389-78	Sequence 78, Appl
37	652	16.9	674	7	US-11-058-727-44	Sequence 44, Appl
38	652	16.9	674	7	US-11-108-389-44	Sequence 44, Appl
39	651.5	16.8	673	7	US-11-058-727-8	Sequence 8, Appl
40	651.5	16.8	673	7	US-11-108-389-8	Sequence 8, Appl
41	651.5	16.8	675	7	US-11-058-727-46	Sequence 46, Appl
42	651.5	16.8	675	7	US-11-108-389-46	Sequence 46, Appl
43	651.5	16.8	677	7	US-11-058-727-84	Sequence 84, Appl
44	651.5	16.8	677	7	US-11-108-389-84	Sequence 84, Appl
45	651	16.8	676	7	US-11-058-727-72	Sequence 72, Appl

ALIGNMENTS

RESULT 1
US-11-058-727-4
; Sequence 4, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad

; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Heirmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnall
; APPLICANT: James P.H. Wong
; APPLICANT: Cao-Guo Yu

; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal

; FILE REFERENCE: 35718/287809

; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15

; PRIOR APPLICATION NUMBER: 60/391,786

; PRIOR FILING DATE: 2002-06-26

; PRIOR APPLICATION NUMBER: 60/460,787

; PRIOR FILING DATE: 2003-04-04

; PRIOR APPLICATION NUMBER: 10/606,320

; PRIOR FILING DATE: 2003-06-25

; NUMBER OF SEQ ID NOS: 134

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 1210

; TYPE: PRT

; ORGANISM: Bacillus thuringiensis

US-11-058-727-4

Query Match 18.2%; Score 703; DB 7; Length 1210;

Best Local Similarity 27.8%; Fred.No. 2.5e-46;

Matches 216; Conservative 127; Mismatches 292; Indels 142; Gaps 31;

QY 5 NDNNEYIIDSHSPYFPNRSNDS-RYPYTNPNQPLQNTNYKWLNNCOGN-TOYGN 62

Db 4 NNQNEYIIDATPS-----TSVNSNDRYPFANEPTNALQNMDDYKDYLKMSAGNASYPGS 59

QY 63 FETFASADTTAAVSAGTIIVSGTLLAGIGLTSISGPILGIIAISFGTLITVFWPAGEQ 122

Db 60 FEVLVSGQD--AAKAAIDIVGKLLISGLG-----VPPVGFIVSLYTLQIDILWPSGQ- 108

QY 123 DKTWTQIKNGEIPVDFTPLTESIKQLKLTLEGFRQLQSYNTALDDWRKLRKLPAGL 182

Db 109 -KQWEIFMEQVEELINQKIAFYARNKALSELEGLGNNYQLYLTALKEWKE----- 158

183 PPSSALQQAALTKIRPENVNDPIREIPGQLETKTLLPIYAQAANFHLNLLQGA 242
159 NPNGS--RALRDVNRNREILDSLTQYMPSPRVNTFEVPLTVYTOAANLHLLLLKDA 216
243 LADEWNADIHPSQIBENAGTSDYYKLLKENIPKYSNYCANTYRTGLKNLRDEPNKWSI 302
217 FGEWGW-----STTTINNYDRQMKLTAEYSDHCWKVETGLAKUGTSAKQWVD 267
303 FNDYRYMTITLDTISQSLYDIKRYDSIGGIEVKIGKINELTREIYVTEINFDRPLQ 362
268 YNQFREMTLTVLDVALFPNYDTRY-----PNETKA---QLTREYVT-----DPLGAV 314
363 RVQ-----PNLATMEYNLTRASFLEQIFPYTENTNFGN-----RLVGISNR 408
315 NVSSIGSWYDKAPSGFVIESSVIRPP-HVFDYITGLTVYVTSQSRSSISSARYIRHWAGH 373
409 DAPYNSNTITETLYGERTGSPPTKTRPPESYKVSIVTDROQSPVPIQP-----HFIIN 463
374 YHRIFSDNIHKOMYGTQNLHSTSF-DFTNYDIYKTLSDAVALLDIVFPGYTYIFFGMP 432
464 QIELYLNGSSNN---TLKYS-----AGGSLSNYQNTTFFQPRKDCNLDVDPGCS-- 511
433 EVEFFWVQNLNTRKTLKPNVSKDIAG-----TRDSELELPETSQ 476
512 PNFNNYSHLSHF-SLFTYSYVIGLQILDTGVLGTHSSVDRYNAISDKIITMIPA 570
477 PNYESYSHRLCHITSIPATGSTTGL-----VPFWSWTHRSADLINAVHSDKITQIPVVK 530
571 GNNL-----DTNSKVIIEGPGHTGMLVYLSQSG-----RLBITCETPNSTQSYFIRLR 618
531 VSDLAPSIITGPNNTVSGFGTGGGIKIRNGVIIISHMRVKIS-----DINKEYSMRIR 586
619 YATNGAGNTLPNLSITIPGVIIPORLNTFSGTNNYNNLYQDGFYQFPSTVTLPLNR 678
587 YAS--ANNTEFYINPSEENVKS-HAQKTMNRGALTYNKNFYATLPPIKFTITE----- 637
679 NIPF-----IFNRADVSNLSIIDKIEFIPITSSMHQNRKQKLETIQTKINTFFTN 730
638 --PFTILGAIFAEADFLGIEAYIDRIEFIPVDETY---EAEQDLEAAKAVNALFTN 689

RESULT 2

US-11-108-389-4
; Sequence 4, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity

FILE REFERENCE: 357118/291049
CURRENT APPLICATION NUMBER: US/11/108,389
CURRENT FILING DATE: 2005-04-18
PRIOR APPLICATION NUMBER: 60/391,786
PRIOR FILING DATE: 2002-06-26
PRIOR APPLICATION NUMBER: 60/460,787
PRIOR FILING DATE: 2003-04-04
PRIOR APPLICATION NUMBER: 10/606,320
PRIOR FILING DATE: 2003-06-25
NUMBER OF SEQ ID NOS: 134
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 1210
TYPE: PRT
ORGANISM: Bacillus thuringiensis

US-11-108-389-4

Query Match 18.2%; Score 703; DB 7; Length 1210;
Best Local Similarity 27.8%; Pred. No. 2.5e-46;
Matches 216; Conservative 127; Mismatches 292; Indels 142; Gaps 31;

QY 5 NDNNEVHIIDSHTSPPYFNRNSNDS-RYPYTNPNQPLONTNYKEWLNMCQGN-TOYGDN 62
DB 4 NNQNEYEIIDATPS---TSVSNDSNRPFPANEPNTALQNNDYKDYKMSAGNASEYFSGS 59
QY 63 FETPASADTTAAVSAGTIVSGTLGAGGLTISISGPITGIIAGIISFCGLTIVVPAGEQ 122
DB 60 PEVLVSGQD--AKAAIDIVKLSGLG-----VPFVGPIVSLYTQLDILMPSQ-- 108
QY 123 DKTWTFQIKMGIFVDTPLTESIKQLKLTLEGFROILQSINTALDDWRKLKRLQAPGL 182
DB 109 -KQWEIFMGEVEELINQKIAEYARNKALSELEGLGNNYQLYLTALBEWKE----- 158
QY 183 PPSALQQAALTKIRPENVNDPIREIPGQLETKTLLPIYAQAANFHLNLLQGA 242
DB 159 NPNGS--RALRDVNRNREILDSLTQYMPSPRVNTFEVPLTVYTOAANLHLLLLKDA 216
QY 243 LADEWNADIHPSQIBENAGTSDYYKLLKENIPKYSNYCANTYRTGLKNLRDEPNKWSI 302
DB 217 FGEWGW-----STTTINNYDRQMKLTAEYSDHCWKVETGLAKUGTSAKQWVD 267
QY 303 FNDYRYMTITLDTISQSLYDIKRYDSIGGIEVKIGKINELTREIYVTEINFDRPLQ 362
DB 268 YNQFREMTLTVLDVALFPNYDTRY-----PNETKA---QLTREYVT-----DPLGAV 314
QY 363 RVQ-----PNLATMEYNLTRASFLEQIFPYTENTNFGN-----RLVGISNR 408
DB 315 NVSSIGSWYDKAPSGFVIESSVIRPP-HVFDYITGLTVYVTSQSRSSISSARYIRHWAGH 373
QY 409 DAPYNSNTITETLYGERTGSPPTKTRPPESYKVSIVTDROQSPVPIQP-----HFIIN 463
DB 374 YHRIFSDNIHKOMYGTQNLHSTSF-DFTNYDIYKTLSDAVALLDIVFPGYTYIFFGMP 432
QY 464 QIELYLNGSSNN---TLKYS-----AGGSLSNYQNTTFFQPRKDCNLDVDPGCS-- 511
DB 433 EVEFFWVQNLNTRKTLKPNVSKDIAG-----TRDSELELPETSQ 476
QY 512 PNFNNYSHLSHF-SLFTYSYVIGLQILDTGVLGTHSSVDRYNAISDKIITMIPA 570
DB 477 PNYESYSHRLCHITSIPATGSTTGL-----VPFWSWTHRSADLINAVHSDKITQIPVVK 530
QY 571 GNNL-----DTNSKVIIEGPGHTGMLVYLSQSG-----RLBITCETPNSTQSYFIRLR 618
DB 531 VSDLAPSIITGPNNTVSGFGTGGGIKIRNGVIIISHMRVKIS-----DINKEYSMRIR 586
QY 619 YATNGAGNTLPNLSITIPGVIIPORLNTFSGTNNYNNLYQDGFYQFPSTVTLPLNR 678
DB 587 YAS--ANNTEFYINPSEENVKS-HAQKTMNRGALTYNKNFYATLPPIKFTITE----- 637
QY 679 NIPF-----IFNRADVSNLSIIDKIEFIPITSSMHQNRKQKLETIQTKINTFFTN 730
DB 638 --PFTILGAIFAEADFLGIEAYIDRIEFIPVDETY---EAEQDLEAAKAVNALFTN 689

RESULT 3

US-11-091-643-6
; Sequence 6, Application US/11091643
; Publication No. US20050246789A1
; GENERAL INFORMATION:
; APPLICANT: TANAKA, Masao
; APPLICANT: YOKOYAMA, Tomoko
; APPLICANT: AOYAGI, Morichi
; APPLICANT: HASEGAWA, Makoto
; APPLICANT: EHARA, Gaku
; APPLICANT: KIMURA, Masaharu
; APPLICANT: NISHIHASHI, Hideji
; TITLE OF INVENTION: Polypeptide having larvae growth inhibiting or
; TITLE OF INVENTION: insecticidal effect on scarabaeidae insects and

```

; TITLE OF INVENTION:  poynuclotide encoding the same
; FILE REFERENCE:  OPI335
; CURRENT APPLICATION NUMBER:  US/11/091.643
; CURRENT FILING DATE:  2005-03-29
; PRIOR APPLICATION NUMBER:  JP 2001-115754
; PRIOR FILING DATE:  2001-04-13
; PRIOR APPLICATION NUMBER:  JP 2001-203463
; PRIOR FILING DATE:  2001-07-04
; NUMBER OF SEQ ID NOS:  22
; SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1386
; TYPE:  PRT
; ORGANISM:  Bacillus popilliae
US-11-091-643-6

```

[illegible]

```

DB              704 k 704
|
RESULT 4
US-11-058-727-14
; Sequence 14, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 673
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis (truncated)
US-11-058-727-14

```

[illegible]

733 K 733

QY 464 QIELYNGSSNN---TLKYS-----AGGSLSNYQNTFFQPRKDCNVLDPGCS-- 511
Db 433 EVEFFMVNQLNTRKTLKYNPVSKOIIAG-----TRDSELEPPETSQ 476
QY 512 PFNNYSYLHSHP-SLFTYSYVIGLQILDTGVLGTHSSVDRYNAISDKIITMIPA 570
Db 477 PNYESYSHRLCHTISIPATGSTTGL-----VPVFSWTHRSADLINAVHSDKTIQIPVK 530
QY 571 GNNL-----DTNSKVIEGPGHTGNNLVYLSQSG-----RLBITCETPNSTQSYFIRLR 618
Db 531 VSDLAPSITGGPNNTVSGPGFTGGGIKVRNGVISHMRVKIS-----DINKEYSMRIR 586
QY 619 YATNGAGNTLPNLSLTIPGVIGIPQRLNNTFSGTNVNNLQYDGFQFPSTVTLPLNR 678
Db 587 YAS--ANNTEFYINPSEENVKS-HAOKTMNRGEALTYNKFNYATLPPIKFTTTE----- 637
QY 679 NIPF-----IFNRADVSNLSIILDKIEFIPI 704
Db 638 --PFITLGAIFEAEDFLGIEAYIDRIEFIV 666
RESULT 5
US-11-108-389-14
; Sequence 14, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE REFERENCE: 357118/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; PRIOR FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis (truncated)
US-11-108-389-14
Query Match 17.5%; Score 678.5; DB 7; Length 673;
Best Local Similarity 27.8%; Pred. No. 8e-45;
Matches 209; Conservative 123; Mismatches 280; Indels 139; Gaps 30;
QY 5 5 NONNEVEIIDSHTSPFPNRSNDS-RYPYTNPNQPLQNTYKELNMCQGN-TQYGDN 62
Db 4 NNQNEVEIIDATPS-----TSVNSNDRYPFANEPTNALQNM DYKYLKMSAGNASBYPGS 59
QY 63 FETTFASADTIAVSACTIVSGTLLAGIGLTSISGPIIGIIGAILISFGTLITVFPAGEQ 122
Db 60 PEVLVSGQD--AKAAIDIVGKLLSLG-----VPFVGPIVSIVTQIIDLIMPSSGQ- 108
QY 123 DKTVMTQFKMGSEIFVDTPLETESIKQLQTLLEGFRQILQSYNTALDDMRKRLQAPGL 182
Db 109 -KSOWEIFMEQVEELNQIAEYARNKALSELGLGNNYQLYLTAEWKE----- 158
QY 183 PPSALQQAALTUKIRFENVNDHFIREFGQFQETKTLPIYQAANFHLNLLQGA 242
Db 159 NPNGS--RALRDVRNREFILDSLFTQMPFSFRVTNFEVPELTVYTAANLHLLKLDASI 216

QY 243 LADENWADIHPSQIEPNAGTSDDDYVYKLLKENIPKYSNYCANTYRTGLKNLDEPNMKWSI 302
Db 217 FGEWGW-----STTTINNYDRQMKLTAEYSDBCHVKWYETGLAKGTSAKQWVD 267
QY 303 FNDYRRYMTITVLDTISQFSLYDIKRYEDSISGGLGVKIKNELTREIYTTETINPRLPOL 362
Db 268 YNFRREMTITVLDVVALFPNYDTRTY-----PMETKA---OLTREYVT-----DPLGAV 314
QY 363 RVQ-----PNLATMEYNLTRASFKLFSFLQFIPTENTNFGN-----RLVGISNR 408
Db 315 NVSSIGSWYDKAPSGFVIESVIRPP-HVFDYITGLTVYTQSRSSISSARYIRHWAGHQS 373
QY 409 DAPYSNTITETLYGERTGSPTKTIRPFESYKYSIVTDROSPPVSPQIP-----HFIIN 463
Db 374 YHRIFSDNIITKMYGTNQNLHSTF-DFTNYDIYKTLSDAVLLDIVFGPGVYIFFGMP 432
QY 464 QIELYNGSSNN---TLKYS-----AGGSLSNYQNTFFQPRKDCNVLDPGCS-- 511
Db 433 EVEFFMVNQLNTRKTLKYNPVSKOIIAG-----TRDSELEPPETSQ 476
QY 512 PFNNYSYLHSHP-SLFTYSYVIGLQILDTGVLGTHSSVDRYNAISDKIITMIPA 570
Db 477 PNYESYSHRLCHTISIPATGSTTGL-----VPVFSWTHRSADLINAVHSDKTIQIPVK 530
QY 571 GNNL-----DTNSKVIEGPGHTGNNLVYLSQSG-----RLBITCETPNSTQSYFIRLR 618
Db 531 VSDLAPSITGGPNNTVSGPGFTGGGIKVRNGVISHMRVKIS-----DINKEYSMRIR 586
QY 619 YATNGAGNTLPNLSLTIPGVIGIPQRLNNTFSGTNVNNLQYDGFQFPSTVTLPLNR 678
Db 587 YAS--ANNTEFYINPSEENVKS-HAOKTMNRGEALTYNKFNYATLPPIKFTTTE----- 637
QY 679 NIPF-----IFNRADVSNLSIILDKIEFIPI 704
Db 638 --PFITLGAIFEAEDFLGIEAYIDRIEFIV 666
RESULT 6
US-11-058-727-2
; Sequence 2, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE REFERENCE: 357118/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1206
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-11-058-727-2
Query Match 17.4%; Score 675; DB 7; Length 1206;
Best Local Similarity 28.2%; Pred. No. 3.5e-44;

Matches	216;	Conservative	125;	Mismatches	300;	Indels	126;	Gaps	32;																																																																																														
Qy	5	NNNNEYIIDSHTSPYFPNNRNSDS--RYPYTNNPNQLONTNTYKEWLNMCOGN--TQYGDN	62	Db	4	NNQNEYEIIDATPS-----TSVSDNSNRYPFANEPTNALQNNDYDKYLKKSAGNASAYPGS	59	Qy	63	PETPASADTTAAVSAGTIVSGTLIAGIGGLTSISGPIGIIICAIISFGTLITVPWPAGEQ	122	Db	60	PEVLVSGD--AAKAAADIVGKLSSGLG-----VPFVGPIVSLVTQLIDILWPSGE-	108	Qy	123	DKTVTTOFIKWGEIFVDTPLTSTIKQLKTLEGFRQILOSYNTALDDWRKKLQAPGL	182	Db	109	-KSQWEIFMEQVEELINQKIAEYARNKALSLEGLGNNYQLYLTALSEEWEE-----	158	Qy	183	PPSALOQAALTKLIRPENVHDFIREIPGOLETYKTLLLPIYAQAANFHNLLOOGAE	242	Db	159	NPNGS--RALRDVRNRFEILDSTQTQMPSRVNTEVPFTVTYAMAANHLHLLKDAISI	216	Qy	243	LADENNADIHPQSIEPNAGTSDDDYYKLLKENIPKYSNYCANTYRTGLKNLRDEPNMKWSI	302	Db	217	FGEWGW-----STTINNYYDRQMKLTAEYSDHCWKWYETGLAKLGTSAKQWD	267	Qy	303	FNDVRRYMTITVLDTISQFSLYDIKRYRDSJGGIEVKGIKNELTREITYTTEINFRLPQL	362	Db	268	YNQPRREWTALAVALFPNDYTRTY-----PMETKA---QLTREYVT-----DPLGAV	314	Qy	363	RVO-----PNLATWEYNLTRASFKLPSFLSQPIFYENTNFNG-RLV-----GIS	406	Db	315	NVSSIGSWYDKAPSFGVIESVIRPP-HVFYDIITGLTVTQSRSISSARYIRHWAGHQIS	373	Qy	407	NRDAPTYSNTITETLYGERTGSPTKIIRPESYKVSVITDRQPPVSPQP-----HPI	461	Db	374	YHRVRSRGSN--LOQMGTQONMLHSTSTP-DEFNTDYIKTLSKDAVLDDIVPGYTYIFFG	430	Qy	462	INOUELYLNGSSNN--TLKYSAGGSLSNYQNTTFQFFPRKDCNCLVIDPCGSCSPNNYS	518	Db	431	MPEVEFFWNQNLNTRKTLKY-----NPVSKDIIASTROSEULELPETSDDQNYSESYS	483	Qy	519	HILSHF-SLFETYSYVIGIQLIQDLTGVLGWHSSHVDRYNAISDKXIITWIPAUK-GNNLD	576	Db	484	HRLCHITSIPATGNTTGL-----VPFESWTHRSADLNNTIYSDKITQIPAVKCWDLNPF	537	Qy	577	NSKVIEGPGHTGNNLV-YLOSQG-----RLBITCTPNSTOSYFIRLYATNGAGNT	627	Db	538	-VPVVKGPGHTGGDLLQNRSTGVSGLTLFARYGLE---KAGKYRVRLRYATA-----	589	Qy	628	LPNISLTIPGVIGHIPPORLNNFTSGTNVNNLOYGDGFQFPSPSTVTLPLNKRNIPIFNRA	687	Db	590	--DIVLHVNDAAIQMPTKMPG-----EDLTSKTFKVADAITTLNLATDSSLALKHNLG	641	Qy	688	DVSNS-----ILLIDKIEFIPITSSHHQNRKQKLETTQTKINTFTTN	730	Db	642	EDPNSTLSGVIYDVRIEFIPVDETY--EAQDLEAAKAWALFTN	685

RESULT 7

```

US-11-108-389-2
; Sequence 2, Application US/11108389
; Publication No. US20050261186A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108.389

```

```

; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1206
; TYPE: Prt
; ORGANISM: Bacillus thuringiensis
US-11-108-389-2

```

Query Match	17.4%;	Score 675;	DB 7;	Length 1206;
Best Local Similarity	28.2%;	Pred. No. 3.5e-44;		

Qy	5	NDNNEYIIDSHTSPYFPNRSND\$-RYPYTNPNQPLONTNYKEWLNMQGN-TOYGDN	62
Db	4	NNQNEYIIDATPS-----TSVNSDN\$RYPANEPTNALQNDYKDYLKM\$AGNASEPGS	59
Qy	63	PEYFASADTTAAVSAGTIIVSGTLLAGIGLTSISPGIIGIALLISGTLITVFWPAGEQ	122
Db	60	PEVLVSQD--AAKAAIDIVGKLLSGIG-----VPFVGPIVSLYTQILDILWPSGE-	108
Qy	123	DKTIVTQFIKMGHIFVDVTPLTESI\$KQLKLTQ\$GPRQILQSYNTALDDWRK\$KLQAPGL	182
Db	109	-K\$QWEIFMEQ\$VEELINQK\$AEARNKAL\$ELEG\$GNNYQ\$LYTAL\$EWE\$-----	158
Qy	183	PPSSALQOAA\$LT\$KIRFENVHNF\$FIREI\$PGFQ\$LET\$YK\$TLL\$PI\$YAOAN\$FHLN\$LQOAG\$	242
Db	159	NPNGS--RALRDVRN\$PEILD\$SLF\$TQMP\$SRV\$NPEV\$PFL\$TVYAMAAN\$HL\$LL\$KDA\$SI	216
Qy	243	LAD\$WNAD\$IHPSQIEPNAGT\$DDY\$YK\$LLKENI\$PKYSNYCANTYRT\$G\$LN\$LRDEPN\$MKWSI	302
Db	217	FG\$EWG\$-----STTINNY\$DROM\$KLTAEY\$SDHC\$VW\$ET\$G\$LAK\$GTS\$AKQ\$WD	267
Qy	303	FNDYRRYMTITVLDTISQF\$LYDI\$KRYRDSIGGIEVK\$IKNELTREIYTT\$EIN\$FDR\$LPQ	362
Db	268	YNQ\$FR\$EMT\$LA\$VDV\$VAL\$PNYDRTY-----PMETKA---Q\$LTREYVT-----D\$PLGAV	314
Qy	363	RVQ-----PNLATMEYN\$LR\$TASF\$K\$LF\$E\$Q\$FIF\$Y\$T\$EN\$N\$FGN-R\$LV-----G\$IS	406
Db	315	NVSSIG\$WYDKAP\$SFGV\$IESSVIRPP-HV\$PDYIT\$GLTVY\$TQ\$RSISSAR\$YIRHWAGH\$QIS	373
Qy	407	NRDAPYSNTIETL\$GERT\$G\$T\$TKTIR\$F\$E\$S\$K\$YSIV\$TDROS\$PP\$V\$PQ-----H\$FI	461
Db	374	YHRSVRGSN--LQOMYGT\$NQLH\$T\$STF-DFTNYDIY\$K\$LSKDA\$V\$LDIVY\$PGYTI\$FFG	430
Qy	462	INQIELY\$NGSSN---TLK\$Y\$AGGS\$LSNYQNTTFFQ\$PRKDCN\$LVID\$PC\$SPN\$FNYS	518
Db	431	MP\$EVE\$F\$M\$NQLN\$NRTK\$LY-----NPV\$KDI\$AST\$D\$SELE\$P\$P\$ET\$DQ\$N\$YESYS	483
Qy	519	HIL\$SHF-SL\$PTYS\$VIGLQ\$ILDT\$GVL\$G\$W\$TH\$SS\$VD\$RYNAISDKIIT\$MI\$PAIK-GNNLDT	576
Db	484	HR\$LCHITSIPATGNTYGL-----VPV\$F\$W\$TH\$S\$ADLN\$NTIY\$SDKITQI\$F\$AVK\$C\$W\$DL\$P\$F	537
Qy	577	NSKVI\$B\$PGHGT\$GNLV-YLQ\$SQ-----RLEITCETPN\$STQ\$SYR\$IR\$LYAT\$NGAGNT	627
Db	538	-VPV\$K\$GPGHGT\$GDL\$QYNR\$STG\$SVG\$TL\$F\$L\$ARY\$LALE---KAGKYR\$V\$R\$LYA\$TDA----	589
Qy	628	LPN\$ISLTI\$PGVIGIPQ\$R\$LAN\$T\$SG\$NYNN\$LYQ\$D\$FGY\$F\$OP\$P\$STV\$T\$PL\$NRN\$T\$P\$FI	687
Db	590	--DIVL\$VND\$AQ\$QMPK\$TWN\$P\$-----EDL\$T\$KT\$FKVAD\$AIT\$LN\$LAT\$D\$SS\$LA\$K\$H\$NLG	641
Qy	688	DV\$SNS-----ILLIDK\$T\$EPI\$IT\$SMH\$QNR\$E\$K\$LETIQT\$KIN\$T\$P\$PTN	730
Db	642	EDPN\$T\$LSG\$VY\$VDR\$IE\$F\$VDETY---EAEQ\$LEA\$AK\$V\$AL\$FTN	695

RESULT 8


```
US-11-091-643-4
; Sequence 4, Application US/11091643
; Publication No. US20050246789A1
; GENERAL INFORMATION:
; APPLICANT: TANAKA, Masao
; APPLICANT: YOKOYAMA, Tomoko
; APPLICANT: AOYAGI, Morichi
; APPLICANT: HASEGAWA, Makoto
; APPLICANT: EHARA, Gaku
; APPLICANT: KIMURA, Masaharu
; APPLICANT: NISHIHASHI, Hideji
; TITLE OF INVENTION: Polypeptide having larvae growth inhibiting or
; TITLE OF INVENTION: insecticidal effect on scarabaeidae insects and
; TITLE OF INVENTION: polynucleotide encoding the same
; FILE REFERENCE: OPI335
; CURRENT APPLICATION NUMBER: US/11/091,643
; PRIOR FILING DATE: 2005-03-29
; PRIOR APPLICATION NUMBER: JP 2001-115754
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: JP 2001-203463
; PRIOR FILING DATE: 2001-07-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1316
; TYPE: PRT
; ORGANISM: Bacillus popilliae
US-11-091-643-4

Query Match 17.3%; Score 669; DB 7; Length 1316;
Best Local Similarity 29.8%; Pred. No. 1.2e-43;
Matches 230; Conservative 106; Mismatches 301; Indels 136; Gaps 34;

QY 1 MNQ-----NNDN-----NEVEII--DSHTSPYPPNNSNDSRYPTNNPNQPLONTNYKE 48
Db 19 MNQYHNDKNSYNQSGNEWIIQPSNALLYSFN-----KPYATDPNVIAEGSYKN 72
QY 49 WLMNCOGNTQYDGNFETASADTIAAVSAGTVSGTLGAGIGLTSIGSPIGI-----IG 103
Db 73 WLDNCTGT--GDTRSPETAISKGAVSAAITIS-----TGLGLGLGVFPASQIG 119
QY 104 AIIISFGTLITVFWPAGEQDKTWTQPIKMGEIFVDTPLETESIKQLKLTQLEGFROILOS 163
Db 120 AF---YTFLLNTLWPA---SNTQWQEIFARVEELINAKLTGHRNSALTKLNGLRNNEI 173
QY 164 YNTALDDWKLKQLQAPGLPPSSALQQAALTAKIRENVHNDFIRETPGFQLETYKTL 223
Db 174 YNEALIVKQ-----DPNNSKLD---DVRSKFVGLNSQFEYIPOPKBEGFEVQL 222
QY 224 PIYAQAANFHLNLLQQAELADENADIIHPSQIEPNAGTSDDYYKLLKENIPKSYNCAN 283
Db 223 TIYAQSANLHLLLRDSSLYGASWGF-----AQATIDNNYNQIRKTAEYANHCTT 273
QY 284 TYRTGLKNIURDENMKWSIFNDYRRYMTITVLTISQFSLYDIKRYRDSIGGIEVKIKN 343
Db 274 WYQTLQRLQGTASSWLSYHRFRREMTITVLDICALFSNYDARSY-----PLEVRG--- 325
QY 344 ELTREIYTYTEIN-----FRLPOLRVQPNLA-TMEYNLTASPKLSPLEQFIPTENTN 397
Db 326 ELTREIYDPVAGTWNIDRAPSFABIELNIVRAPRTVWISGDLIVYTGRLYGYTGND 385
QY 398 F--GNRLVGISNR---DAPTYNTITETLYGERTGSPTKTIRPFESYKVSIVTDROSP 451
Db 386 YNAHRLDLETNGYRPEGTGYSTINIS-----RTUSIPWNSI---DVSSTVTVTGSAA 438
QY 452 PVSPIQPHFIINQIYELYNLNGSNNTLYKSAGGSLSN--YQNTTFQFPKCKCNLVID-P 508
Db 439 PTG----GFVLGVASARFFSKSPST-----GLGERVQNPVYF-----SSSTLTFLNP 483
QY 509 GC---SPNFNNYSHILSHSLFTSYVIGLQILDTGVLGWTWSSVDYRNALSKLIITM 565
Db 484 GVDQDPTTAADYSKLSKSCITAFR---TGINGTV---PVFGYSATVSRDNRNIEPKITQ 536
```

```
566 IPAIGKNLDTNSKVIIEGPGHTGGNLVYLSQGRLEITCTETPNSTQSYFIRLVAT----- 621
537 IPAVKNSLD--NCPVVRGTGFTGGDWLTKTSYLSVFLTITSSRAGQSYRIRVYAAAADL 595
622 -NGAGNTLPNLSLTIPGVIGIPQRLANTFSGTNNYNLOYGDFGVFPFPSTVTLPLNRI 680
596 IMSIYNDPFIISKISLTKSMPP--LTET---VPYFAKFPADFGV-----TITATANKRY 646
681 PFIFNRADVSNSILIIDKIEFIPITSSMHQNRKQKLETTIQTKINTFTFTHTK 733
647 TFOFH---TGGAALIDRIEFVPIEGSLFEYETKQLEKAKAVNHLFTDGSK 695

RESULT 9
US-11-058-727-74
; Sequence 74, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnall
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis (mutated)
US-11-058-727-74

Query Match 17.0%; Score 657.5; DB 7; Length 675;
Best Local Similarity 28.5%; Pred. No. 3.3e-43;
Matches 211; Conservative 120; Mismatches 293; Indels 117; Gaps 31;

QY 5 NDNNEYIIDSHTSPYPPNNSNDS--RYPYTNPNQPLQNTYKELNMCQGN--TOYGDN 62
Db 4 NNQNEVEIDATFS-----TSVSNDSNRYPANFPTNALQNDYKDYKMSGNAASEVPGS 59
QY 63 FETFASADTIAAVSAGTVSGTLGAGIGLTSIGSPIGIIGAILISFGTLITVFWPAGEQ 122
Db 60 PEVLVSQD--AAKAAIDIVGLLSGLG-----VPFVGPIVSLYTLQIDILWPSC-- 108
QY 123 DKTWTQPIKMGEIFVDTPLETESIKQLKLTQLEGFROILOSNTALDDWKLKQLQAPGL 182
Db 109 -KSQWEIFMEQVEELINQKIAEYARNKALSELEGNNYQLYLTALBEWEE--NPPRSRGF 166
QY 183 PPSALQQAALTAKIRENVHNDFIRETPGFQLETYKTLPIYAQAANFHLNLLQQAEL 242
Db 167 RSGALRD---VRNRFELDSLFTQYMPFRVTNFVFPFLTVYMAANLHLLLKDAI 222
QY 243 LADEWNADIIHPSQIEPNAGTSDDYYKLLKENIPKSYNCANTYRTGLKNIURDENMKWSI 302
Db 223 FGEWGN-----STTINNYDRQMKLTAEYSCHVCVKWYETGLAKLKGTSAKQWVD 273
QY 303 FNDYRRYMTITVLTISQFSLYDIKRYRDSIGGIEVKIKNELTREIYTTTINFDRLPOL 362
Db 274 YNQFRREMTILAVLDVVALFPNYDTRY-----PMETKA---QLTREVYTT-----DPLGAV 320
```



```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis (mutated)
US-11-108-389-74

Query Match      17.0%; Score 657.5; DB 7; Length 675;
Best Local Similarity 28.5%; Pred. No. 3.3e-43;
Matches 211; Conservative 120; Mismatches 293; Indels 117; Gaps 31;

QY 5 NDNVEYIIDSHTSPYFPNRSNDS--RYPYNNPNQPLONTNYKEWLNMCQGN--TQYGDN 62
Db 4 NNQNEYIIDATPS-----TSVNSNDRYPFANEPTNALQNM DYKDYKMSAGNASEYPGS 59
QY 63 FETFASADTTAAVSAGTIVSGTLLAGIGLTSISGPIIGAIISFGTLITVFWPAGEQ 122
Db 60 PEVLVSGQD--AKAAIDIVGKLLSGLG-----VPFVGPIVSLYTLQILDILWPSGE- 108
QY 123 DKTWTQFIMGGEIFVDTPITESIKQLKQLTLEGFRQILQSINTALDDMRKLRQAAPGL 182
Db 109 -KSQWEIFMEQVEELNQIAEYARNKALSELEGNNYQLYLTALAEWEE--NPFERSRGF 166
QY 183 PPSALQQAALTUKIRPENVHNDPIREIPGFOLEYTKTLPIYAQAANFHLNLQOGAE 242
Db 167 RSGALRD-----VRNRFELDSLFTQMPFSFRVTNFEVPEFLTYYAMAANLHLLKDAI 222
QY 243 LADEWNADIHPSQIEPNAGTSDDYKLLKENIPKYSNYCANTYRTGLKLRDEPNMKWSI 302
Db 223 FGEWGW-----STTINNYDRQMKLTAEYSDHCWKVETGLAKUGTSAKQWVD 273
QY 303 FNDYRYMTITVLDITISQFSLYDIKRYRDSIGGIEVKIGIKNELTREIYTTTEINFDRLPQL 362
Db 274 YNQFREMTLAVLDVVALFPNYDTRY-----PMETKA---QLTREYVT-----DPLGAV 320
QY 363 RVQ-----PNLATMEYNLTRASPKLFSFLEQIFFYTENTNFGN--RLV-----GIS 406
Db 321 NVSSIGSWYDKAPSGFVIESSVIRPP--HVFDTYITGLTVYTSQSRISISSARYIRHWAGHQS 379
QY 407 NRDAPTYNTITELTGERTGSPTTKIRPPESYKYSIVTDROSPVSIQPIQ-----HFI 461
Db 380 YHRVSRGSN--LQOYGTQNQLHSTSTF--DFTNYDIYKLSKDAVLDDVYPGYTIFFG 436
QY 462 INQIELYLGSSNN---TLKYSAGGSLSNYQNTFFQFPRKDCNVLVIDPGCSNPNFNYS 518
Db 437 MPEVEFFMVLNQLNTRKTLKY-----NPVSKDIIASTRDSLELPETSQDQNYESYS 489
QY 519 HILSHF--SLFTSYVIGLQQLDGTGLVGLWTHSSVDRYNAISDKIITMIPAIC--GNLDT 576
Db 490 HRLCHITSIPATGNTTGL-----VPVFSWTHRSADLNNTIYSDKITQIPAVKWCNDLPP 543
QY 577 NSKVIEGPGHTGNLV--YLQSQ-----RLEITCETPNSTQSYFIRLYATNGAGNT 627
Db 544 -VPVVGPGHTGDDLQYNRSTGVSGLTFLARYGLALE---KAGKYRVLRYATDA---- 595
QY 628 LPNISLTI PGVIGIPQRLNNTSGTNYNNLQYDGFYQFPGTPTVTLPLNRNIPFIFNRA 687
Db 596 --DIVLHVDAQIQMPTWNP-----EDLTSKTFKVAADAITVNLATDSSVAVKHNLG 647
QY 688 DVSNS----ILIIDKIEFPI 704
Db 648 EDPNSTLSGIVYVDRIEFIPV 668

RESULT 12
US-11-108-389-80
; Sequence 80, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
```

Db 596 --DIVLHVNDAAQIQMPKTNPG-----EDLTSKTFKVADAITVNLATDSSVAVKHNLG 647

Qy 688 DVSN-----ILIIDKIEFPI 704

Db 648 EDPNSTLSGIVYVDRIEFIPV 668

RESULT 13

US-11-058-727-82

; Sequence 82, Application US/11058727

; Publication No. US20050261483A1

; GENERAL INFORMATION:

; APPLICANT: Andre R. Abad

; APPLICANT: Ronald D. Flannagan

; APPLICANT: Rafael Herrmann

; APPLICANT: Theodore W. Kahn

; APPLICANT: Albert L. Lu

; APPLICANT: Billy Fred McCutchen

; APPLICANT: James K. Presnail

; APPLICANT: James F.H. Wong

; APPLICANT: Cao-Guo Yu

; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal

; FILE REFERENCE: 35718/287809

; CURRENT APPLICATION NUMBER: US/11/058,727

; CURRENT FILING DATE: 2005-02-15

; PRIOR APPLICATION NUMBER: 60/391,786

; PRIOR FILING DATE: 2002-06-26

; PRIOR APPLICATION NUMBER: 60/460,787

; PRIOR FILING DATE: 2003-04-04

; PRIOR APPLICATION NUMBER: 10/606,320

; PRIOR FILING DATE: 2003-06-25

; NUMBER OF SEQ ID NOS: 134

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 82

; LENGTH: 674

; TYPE: PRT

; ORGANISM: Bacillus thuringiensis (mutated)

US-11-058-727-82

Query Match 17.0%; Score 657; DB 7; Length 674;

Best Local Similarity 28.5%; Pred. No. 3.6e-43;

Matches 212; Conservative 118; Mismatches 291; Indels 122; Gaps 31;

Qy 5 NDNNEYIIDSHTSPYFPNRSNDS-RYPYTNPNQPLQNTNYKWLNMCOGN-TOYGDN 62

Db 4 NNQNEYIIDATPS-----TSVNSNDRYPFANEPTNALQNDYKYDKLXMSAGNASEYPGS 59

Qy 63 FETPASADTTAAVSAGTIVSGTLGAGLTSISGPIGIIIGAIISFGTLITVFWPAGEQ 122

Db 60 PEVLVSGQD--AAKAAIDIVGKLISGLG-----VPFVGPIVSLYTLQILDILWPSGE- 108

Qy 123 DKTWVTOFKMGEIFVDTPLTSTESIKQLKQLTLEGFRQILQSYNTALDDWRKLRQAPGL 182

Db 109 -KSQWEIFMEQVEELINOKIAEYARNKALSELEGNNYQLYLTALBEEWE-----N 159

Qy 183 PPSSAL--QQAALTKEFEENVHDFEIPGFOLEYTKLLPIYAAANFHLNLQOG 240

Db 160 PNGSRFRQALRDVRNRFELDSLFTQYMPFRVTNFEVPLFVYAAANLHLLLLKDA 219

Qy 241 AELADEWNADIHPQIBFNAGTSDYYKLLKENIPKYSNYCANTYRTGLKMLRDEPNMKW 300

Db 220 SIFGEENGW-----SITTIINNYDRQMKLTAEYSDHCVKWYETGLAKLKGTSAKQW 270

Qy 301 SIFNDYRYMTITVLTISQPSLYDIKRYRDSIGGI EVKGIKBLTREIYTTENFRRLP 360

Db 271 VDNQPRREMTLAVLDVALFPNYDTRTY-----PMETKA---QLTREYVT-----DFLG 317

Qy 361 QLRVQ-----PNLATMEYNLTASFKLFSLEQEIFYTTENTNFGN-PLV-----G 404

Db 318 AVNVSSIGSWYDKAPSGVIESSVIRPP-HVFDYITGLTVYTSQRSSSARYIRHWAGH 376

Qy 405 ISNRDAPYSNTITETLYGERTGSPPTKTIREFSKYSIVTDROSPPVSIQOP-----H 459

Db 377 ISYHRVSRGSN--LQOMYGTQNQLHSTSTP-DFTNYDIYKTLKSKOAVLDDIVPGTYIF 433

Qy 460 FIINQIELYLNAGSSNN---TLKYSAGGSLSNYQNTTFFQPPRRKOCNLDVIDPGCSNPNFN 516

Db 434 FGMPEVEFFMVNQLNNTRKTLKY-----NPVSKDIIASTRDSLELEPETSDDQNYES 486

Qy 517 YSHLSHP-SLPYTSYVIGLQILDTGVLGTHSHSVDRYNAISDKIIMIPAIK-GNNL 574

Db 487 YSHRLCHTITSIPATGNTTGL-----VPVFSWTHRSADLNNTIYSDKITQIPAVKCDNL 540

Qy 575 DTNSKVIKSGHGTGNLV-YLQSQG-----RLEICETPNSQTSYFIRLRYATNGAG 625

Db 541 PF-VPVVGPGHGTGDLLOYNRSTGVTGLFARYGLALE---KAGKYRVRURYATDA-- 594

Qy 626 NTLNPNISLTIPGVIGIPQRLNNTFSGTNNYNLQYGFQFPSPSTVTLPLNRNIPFIEN 685

Db 595 ----DIVLHVNDAAQIQMPKTNPG-----EDLTSKTFKVADAITVNLATDSSVAVKEN 644

Qy 686 RADVSNS----ILIIDKIEFPI 704

Db 645 LGEDPNSTLSGIVYVDRIEFIPV 667

RESULT 14

US-11-108-389-82

; Sequence 82, Application US/11108389

; Publication No. US20050261188A1

; GENERAL INFORMATION:

; APPLICANT: Andre R. Abad

; APPLICANT: Ronald D. Flannagan

; APPLICANT: Rafael Herrmann

; APPLICANT: Theodore W. Kahn

; APPLICANT: Albert L. Lu

; APPLICANT: Billy Fred McCutchen

; APPLICANT: James K. Presnail

; APPLICANT: James F.H. Wong

; APPLICANT: Cao-Guo Yu

; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal

; FILE REFERENCE: 35718/291049

; CURRENT APPLICATION NUMBER: US/11/108,389

; CURRENT FILING DATE: 2005-04-18

; PRIOR APPLICATION NUMBER: 60/391,786

; PRIOR FILING DATE: 2002-06-26

; PRIOR APPLICATION NUMBER: 60/460,787

; PRIOR FILING DATE: 2003-04-04

; PRIOR APPLICATION NUMBER: 10/606,320

; PRIOR FILING DATE: 2003-06-25

; NUMBER OF SEQ ID NOS: 134

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 82

; LENGTH: 674

; TYPE: PRT

; ORGANISM: Bacillus thuringiensis (mutated)

US-11-108-389-82

Query Match 17.0%; Score 657; DB 7; Length 674;

Best Local Similarity 28.5%; Pred. No. 3.6e-43;

Matches 212; Conservative 118; Mismatches 291; Indels 122; Gaps 31;

Qy 5 NDNNEYIIDSHTSPYFPNRSNDS-RYPYTNPNQPLQNTNYKWLNMCOGN-TOYGDN 62

Db 4 NNQNEYIIDATPS-----TSVNSNDRYPFANEPTNALQNDYKYDKLXMSAGNASEYPGS 59

Qy 63 FETPASADTTAAVSAGTIVSGTLGAGLTSISGPIGIIIGAIISFGTLITVFWPAGEQ 122

Db 60 PEVLVSGQD--AAKAAIDIVGKLISGLG-----VPFVGPIVSLYTLQILDILWPSGE- 108

Qy 123 DKTWVTOFKMGEIFVDTPLTSTESIKQLKQLTLEGFRQILQSYNTALDDWRKLRQAPGL 182

Db 109 -KSQWEIFMEQVEELINOKIAEYARNKALSELEGNNYQLYLTALBEEWE-----N 159

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 11, 2005, 22:52:13 ; Search time 9651 Seconds

(without alignments)
4329.079 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869

Sequence: 1 MNQNDNNEVEIIDSHSPV.....KLETIQKINTFTNHTKTL 735

Scoring table: BLOSSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 2842172563 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10783417/runat_01122005_141956_24331/app.query.fasta_1.903
-DB=GenEmbl -OFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blossum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10783417 @CGN 1 1 6416 @runat_01122005_141956_24331 -NCPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DE TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sts.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3869	100.0	2208	6	CQ868305 Sequence
2	3314	85.7	2235	6	CQ868307 Sequence
3	3104.5	80.2	2085	6	CQ868309 Sequence

	4	1061.5	27.4	3543	1	BTTOXD1
5	1061.5	27.4	4451	6	I08083	Y00423 Bacillus th
6	1061.5	27.4	4451	6	I09103	I08083 Sequence 1
7	1060.5	27.4	4253	6	BACISRH4	I09103 Sequence 1
8	1060.5	27.4	4934	6	E01676	D00248 Bacillus th
9	1060.5	27.4	127923	1	BTBPTOXIS	E01676 DNA sequenc
10	897.5	23.2	3753	1	BACMSQB	AL731825 Bacillus
11	897.5	23.2	127923	1	BTBPTOXIS	ML2662 B.thuringie
12	880	22.7	3756	6	E00614	AL731825 Bacillus
13	862.5	22.3	4959	1	AF285775	E00614 DNA encodin
14	850	22.0	6009	1	BTB251977	AF285775 Bacillus
15	849	21.9	4896	1	AF132928	AJ251977 Bacillus
16	808.5	20.9	3752	1	AB125059	AF132928 Bacillus
17	784	20.3	3644	1	BTB251978	AB125059 Bacillus
18	775.5	20.0	3507	1	BTU04365	AJ251978 Bacillus
19	775.5	20.0	3507	6	I25972	U004365 Bacillus th
20	759	19.6	3684	1	BTTOXD2	I25972 Sequence 3
21	759	19.6	3684	6	I08884	X07423 Bacillus th
22	757.5	19.6	3535	6	I06096	I08884 Sequence 2
23	754	19.5	3684	1	BTITOX	I06096 Sequence 2
24	752.5	19.4	4186	1	BACISRH3	X07082 Bacillus th
25	752.5	19.4	4186	6	E01905	D00247 Bacillus th
26	746	19.3	3642	1	AB193814	E01905 genomic DNA
27	709	18.3	3668	1	D88381	AB193814 Bacillus
28	705	18.2	4391	1	BTB2665	D88381 Bacillus th
29	705	18.2	6698	1	AB116651	Y07603 B.thuringie
30	703	18.2	3633	6	CS130948	AB116651 Bacillus
31	703	18.2	3633	6	AX543926	CS130948 Sequence
32	703	18.2	6613	6	CS130963	AX543926 Sequence
33	703	18.2	6613	6	AX543951	CS130963 Sequence
34	696.5	18.0	4056	1	BACCRID2	AX543951 Sequence
35	696	18.0	3536	6	E01029	M20242 B.thuringie
36	689.5	17.8	4188	6	BD177494	E01029 DNA sequenc
37	689.5	17.8	4188	6	AX670919	BD177494 Polypepti
38	687	17.8	3931	6	AB161456	AX670919 Sequence
39	678.5	17.5	2003	6	AX543930	AB161456 Bacillus
40	678.5	17.5	2022	6	CS130958	AX543930 Sequence
41	678.5	17.5	2022	6	AX543940	CS130958 Sequence
42	677	17.5	8572	1	AB116649	AX543940 Sequence
43	677	17.5	8605	1	AB116650	AB116649 Bacillus
44	676	17.5	1953	6	I25973	AB116650 Bacillus
45	676	17.5	1977	1	BTCADE	I25973 Sequence 5
						X17123 Bacillus th

ALIGNMENTS

RESULT 1	CQ868305	Sequence 6 from Patent WO2004074462.	2208 bp	DNA	linear	PAT 13-SEP-2004
LOCUS	CQ868305	Sequence 6 from Patent WO2004074462.	2208 bp	DNA	linear	PAT 13-SEP-2004
DEFINITION	CQ868305	Sequence 6 from Patent WO2004074462.	2208 bp	DNA	linear	PAT 13-SEP-2004
ACCESSION	CQ868305	Sequence 6 from Patent WO2004074462.	2208 bp	DNA	linear	PAT 13-SEP-2004
VERSION	CQ868305.1	GI:51998351	2208 bp	DNA	linear	PAT 13-SEP-2004
KEYWORDS	CQ868305.1	GI:51998351	2208 bp	DNA	linear	PAT 13-SEP-2004
SOURCE	CQ868305.1	GI:51998351	2208 bp	DNA	linear	PAT 13-SEP-2004
ORGANISM	CQ868305.1	GI:51998351	2208 bp	DNA	linear	PAT 13-SEP-2004
REFERENCE	CQ868305.1	GI:51998351	2208 bp	DNA	linear	PAT 13-SEP-2004
AUTHORS	CQ868305.1	GI:51998351	2208 bp	DNA	linear	PAT 13-SEP-2004
TITLE	CQ868305.1	GI:51998351	2208 bp	DNA	linear	PAT 13-SEP-2004
JOURNAL	CQ868305.1	GI:51998351	2208 bp	DNA	linear	PAT 13-SEP-2004
FEATURES	CQ868305.1	GI:51998351	2208 bp	DNA	linear	PAT 13-SEP-2004
source	CQ868305.1	GI:51998351	2208 bp	DNA	linear	PAT 13-SEP-2004
CDS	CQ868305.1	GI:51998351	2208 bp	DNA	linear	PAT 13-SEP-2004

/translation="MNQNDNVEYIIDSHTSPYFENRNSDSRYPTNNPNQPLQNT
NYKWLNMCOQNTVQDNFETPASADTIIVAGSAGTIIVSGTLLAGIGLTSISGPIGI
GAILISGTLITVFWAGSODKTWTQVFIKMGSEIFVDTPLETESIKOLKULLEGFRQI
LOSNTALDDWRKRLQAPGLPSSALQQAALTKIRFENVHNDFTREIPGQLETY
KTLILLPIYAAANPHNLNQQOGLADEWNADTHPSQIBNAGTSDDYKLLKENTPK
YSNYCANVTGTGKNIIRDENMKWSIPNDYRVTMTITVLDTI SQFSLYDIKRVDSIG
GIEVKGINKELTREIYTHETINFRDLRQLRVQPNLATMEYNLIRASPKLSFLQPIFY
TENTNFRNLVGLSNDAPYSNTITETLYGERTGSPYTKIIRPFESYKVSIVTDQOS
PPVSPLOPHFINQIELYLNGSNNTLKYSAGGSLSNYQNTTFFQPKKDCNLDVDP
GSPSPNNYTHSLFTSYVIGLOQLIDTLGVLTGWTHSSVDRYNAISDKIITMI
PAIKGNLNDTNSKVIQPGHGTGNNLVLOSQGRLEITCETPNSTQSYFRLRYATNGA
GNTLPNLSITPGVIGIPORLNNTPSGTNNYNLOYDGYFQFQPSVTPLPLNRNIPF
INRADVSNLSILLIDKIEPIPTSSMHQREKQLETIQIKINTFFFNHFKTL"

ORIGIN

Alignment Scores:
Pred. No.: 5.4e-283 Length: 2208
Score: 3869.00 Matches: 735
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-783-417-2 (1-735) x CQ868305 (1-2208)

QY 1 MetAsnGlnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
DB 1 ATGAATCAAAATAACGATAATAACGAATATGAAATATTGATTCGCGCATACCTTAT 60
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
DB 61 TTTCCGAACAGAAACAGTAGTATGATTCGTAGATACCTTTACACAAATAATCAACCA 120
QY 41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
DB 121 TTACAAACACAAATTACAAAGAGTGCTCAATATGTGTCAAGGAATAACAAATATGGT 180
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80
DB 181 GATAATTTCCGAGACATTTGCTAGTGTGATACAAATTTGCTGAGTGTAGTGTACTATT 240
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
DB 241 GTATCCGGTACTCTGTAGCGGTATAGGTGGGCTCACTTCTATATCCGACCGATAGGA 300
QY 101 IleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGly 120
DB 301 ATAAATAGGTGCTATAATAATATCTTTTGGTACCCTTAATCACTGTCTTTTGGCCGCGGGA 360
QY 121 GluGlnAspLysThrValThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
DB 361 GAACAGACAAACAGTAGGACACAAATTTATTAATGGAGAAATTTTGTGTGATACA 420
QY 141 ProLeuThrGluSerLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
DB 421 CCGTTAACAGAAAGCATAAACAGCTAAAGTTACAAACTTTAGAGGATTTAGACAAATA 480
QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaPro 180
DB 481 TTACAAAGCTATTAATAACAGATTTAGATTGGAGAAATTTAAAGAGACTTACAAAGCTCCT 540
QY 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGlu 200
DB 541 GGATTAACCATCATCAGCATTTAGATTGATTTGAGAAATTTAAAGAGACTTACAAAGCTCCT 600
QY 201 AsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThr 220
DB 601 AATGTTTCACAATGATTTTATTCAGAAATACCTGGTGTTCACACTTGAACCTTATAAACG 660
QY 221 LeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGly 240
DB 661 CTATTACTACCTATTATTATGCGCAAGCTGCTGCTATTTTCAATTTAAATTTTATTTACAAACAGGT 720

QY 241 AlaGluLeuAlaAspGluThrProAsnAlaAspIleHisProSerGlnIleGluProAsnAla 260
DB 721 GCTGAATTTGCTGATGAATGAATGCAATATACATCTCTTCCACAAATTTGAACCTAATGCT 780
QY 261 GlyThrSerAspAspTyrTyrLysLeuLysGluAsnIleProLysTyrSerAsnTyr 280
DB 781 GGAACATCAGATGACTATTATAAACTTTTAAAGAAATATACCTAATAATATAGTAACAT 840
QY 281 CysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrp 300
DB 841 TGTGCAAAATACCTATAGAACAGGACTAAAAATCTTAGAGACGACCAAAATATGAATGG 900
QY 301 SerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGln 320
DB 901 AGTATATTTAATGACTATCGAAGATATATGACCAATTAATCTGATTTAGATACCATCTCTCAA 960
QY 321 PheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluValLysGly 340
DB 961 TTTTCTTTTATGATATAAAAGATATAGAGATTTCAATAGGAGGATAGAAATGTAAGGC 1020
QY 341 IleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuPro 360
DB 1021 ATTAAGAATGAACCTCACAGAGAAATTTATACAACTGAATAAATTTTGATCGTCTTCT 1080
QY 361 GlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPhe 380
DB 1081 CAACCTTAGAGTTCAACCCCAATCTAGCTACGATGGAATATAATTTAAACACGTCGAAGTTTT 1140
QY 381 LysLeuPheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsn 400
DB 1141 AAATATATTTTCATTTTATAGAACAAATTTATTTTATACAGAAATACAAATTTTCGGGAAT 1200
QY 401 ArgLeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
DB 1201 CGTTAGTGGTATTTCTAATCGTAGTGACCTACTTATAGCAATATCTATTAACCTGAACT 1260
QY 421 LeuTyrGlyGluArgThrGlySerProThrLysThrIleArgProPheGluSerTyr 440
DB 1261 TTATATGGAGAAAGAACAGGTTCACCCACACAAACAAATAAGACCATTTGAATCTTAT 1320
QY 441 LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
DB 1321 AAAGTTTCAATTTGTAATGATAGACAAATCACCTCTCTCTTCCCTATTTCAACCACTTT 1380
QY 461 IleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsnThrLeuLysTyrSer 480
DB 1381 ATAAATTAATCAATTTGAACCTTTATTAATGGCTCATCTAACACACACTCAATATTTCA 1440
QY 481 AlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArgLysLys 500
DB 1441 GCAGGAGGGTCTTTATCTAATTTACAAACACAACTTTTTCATTTCTCTAGAAAAA 1500
QY 501 AspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIle 520
DB 1501 GACTGCAATCTAGTTATGATCCAGGTGTTCACCAAACTTTTAAATACTATAGTCATATT 1560
QY 521 LeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyGlnLeuGlnIleLeu 540
DB 1561 TTATCCCATTTTTCATTTATTTACTTATTCCTATGTGATTTGATTTACAGCTACAAATATTA 1620
QY 541 AspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerAsp 560
DB 1621 GATACAGGTGTATAGGATGGACACACAGTAGTGTGTGATAGATATATGCAATATCAGAT 1680
QY 561 LysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysVal 580
DB 1681 AAAATAATTAACAATGATCCAGCAATCAAAGGTAAACAATCTTGATACAAACTCTAAGGTA 1740
QY 581 IleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeu 600
DB 1741 ATTTCAAGGACCTGGTTCATACAGGAGGAACTTGGTTTTATTTACAAAGTCAAGGGCGTTTA 1800
QY 601 GluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAla 620

|||||
1801 GAAATTACATGTGAAATCTCTAATTCACACAATCTTATTTCATTAGACTTCGATATGCT 1860
QY 621 ThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerIleuThrIleProGlyValIleGly 640
Db 1861 ACAANTGGTCTCGAAATACTCTTCTTAATATATCTCTTCAATATACAGGAGTAAATAGGA 1920
QY 641 IleProGlnArgLeuAsnAsnThrPheSerGlyThrAsnThrIleAsnThrLeuGlnThr 660
Db 1921 ATACCACCTCAGCAGCTCAACAACACTTTCTCGTACAAATTAATAATTTTAAATACATAC 1980
QY 661 GlyAspPheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIle 680
Db 1981 GGAGATTTTGGGTATTTCCAAATTTCCAAATGACAGTACATACCTTTAAATCGAACAATA 2040
QY 681 PropheIlePheAsnArgAlaAspValSerAsnSerIleleuIleAspLysIleGlu 700
Db 2041 CCATTTATATTAATCGTGAGATGTATCAAAATTCAAATTTTAAATCAATTTGATAAAATTTGAA 2100
QY 701 PheIleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIle 720
Db 2101 TTTATACCAATTAATCTCTCTATGACCAAAATAGAGAAAACAAAATTAGAAAACATATC 2160
QY 721 GlnThrLysIleAsnThrPhePheThrAsnHisThrLysThrLeu 735
Db 2161 CAACAATAATAATACATTTTTCACAAATCATACAAAACACATT 2205

RESULT 2
CQ868307 2235 bp DNA linear PAT 13-SEP-2004
LOCUS Sequence 8 from Patent WO2004074462.
DEFINITION CQ868307
ACCESSION CQ868307.1 GI:51998353
VERSION
KEYWORDS
SOURCE
ORGANISM
Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE
1
AUTHORS Carozzi, N., Hargiss, T., Koziel, M.G., Duck, N.B. and Carr, B.
TITLE Delta-endotoxin genes and methods for their use
JOURNAL Patent: WO 2004074462-A 8 02-SEP-2004;
Athenix Corporation (US)

FEATURES
source
1..2235
/organism="Bacillus thuringiensis"
/mol_type="unassigned DNA"
/db_xref="taxon:1428"
1..2235
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAH33949.1"
/db_xref="GI:51998354"
/translation="MKQNNNEVEIIDSKNLSVPSNRNIDHSRYPYTNPNQBLQNTN
YKELNMQCQNTQYQNFENFASADTAAVSAGTIVSGTLGAGIGTSSGFIGIG
AIIISFGLTIVFWPEQDKWTQFIKMGEIFVDTPLETISKQLQTLLEGFQIL
QSYNTALDDWRKLQAPGLPSSALQQAALTKIRFENVHNDFTREIPGFOLETYK
TLALLPIVAQAAPHNLQQAELADEWNADHPQIENPAGTSDDYKILKENIPKY
SNYCANTYRGLNKLNEPNRWSIENDYRMYTITVLDITIAQFSFVDIKRYKDSIGR
IGGKTELREIYTFTEINFORLTYLEIQPNLAIMENYLRSGRLFSFDELIFYTKN
ETVGNLAVGTANRSTYATGTETIYGERTPPTTKLIPFESYKSVIVTDQVTP
SPPNYITFQINQELYNLNSPSNKLTYSAGNLSNDKKTDFQFPVKDKCKPIINPNC
LPNSYVSHITLSQFSPNYKYKIGLAINILYTGALGWTSSVNRNNAISDKIITMIPA
IKGNSLDTSNKVEGHTCGNLVYLOSQGLEITCETPNSTQSYIRLYRYNAGN
TLNLSITLPGVIGIPQRLNNTFSGTYNNLQIGDFYQFSPFTVPLNRPPIIF
NRADVSNLIIIDIKFIPITPSTSSVRQNRKQKLETTQTAKINTFTTNHNTNLINATN
YDID"

ORIGIN

Alignment Scores: 5.39e-241 Length: 2235
Score: 3314.00 Matches: 638

Percent Similarity: 91.28% Conservative: 32
Best Local Similarity: 86.92% Mismatches: 60
Query Match: 85.66% Indels: 4
DB: 6 Gaps: 3
US-10-783-417-2 (1-735) x CQ868307 (1-2235)
QY 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 1 GTGAATCAAAATAAT--AAATAATGAATATGAGATATTCGATTCGAAGAATTTATCTTAT 57
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 58 CTTTCTAACAGAAATATGATCATTTCTAGATACCCTTACACAAATAATCCAAATCAACCA 117
QY 41 LeuGlnAsnThrAsnTyrLysGluTyrIleuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
Db 118 TTCAAAACACAAATTAACAAGAGTGGCTCAATATGTCTCAAGGGAATACACAATATGCT 177
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaIleValSerAlaGlyThrIle 80
Db 178 GATAAATTTCCAGACATTTGCTAGTGTGATACAAATGCTGAGTTAGTGCAGGTACTATT 237
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
Db 238 GTATCCGGTACTCTGTAGCCGTATAGTGGGCTCACTTCTATATCCGACCGATAGGA 297
QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120
Db 298 ATAATAGTGTCTATAATAATATCTTTTGGTACCCTTAATCACTGTCTTTTGGCCCGGGA 357
QY 121 GluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db 358 GAACAAGACAAAACAGTATGACACAAATTTTAAATGGGAGAAATTTTGTGTGATACA 417
QY 141 ProLeuThrGluSerIleLysGlnLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
Db 418 CCGTTAACAGAAAGCATAAACACAGCTAAAGTTTACAACTTTTAGAAGGATTTAGACAAATA 477
QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaPro 180
Db 478 TTCAAAAGCTATATAATACAGCATTTAGATGATGGAGAAAATTAATAAGACTCAAGCTCT 537
QY 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGlu 200
Db 538 GGATTACCACCATCATCAGCATTAACAACAGCTGCTTGAATCTTTAAATAATACGATTTGAG 597
QY 201 AsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThr 220
Db 598 AATGTTCCACAATGATTTTATTCGAGAAATACCTGGTTTCCAACTTGAACACTTATAAAACG 657
QY 221 LeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGly 240
Db 658 CTATTACTACTATTATTCGCAAGCTGCTAATTTTCAATTAATTTATTAACAACAGGT 717
QY 241 AlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAla 260
Db 718 GCTGAATGGCTGATGAATGGAATGCAGATATACATCCTTCACAAAATTTGAACCTAATGCT 777
QY 261 GlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyr 280
Db 778 GGAACATCAGATGATTAATAAACTTTTAAAGAAAATATACCTTAAATATATAGTAACAT 837
QY 281 CysAlaAsnThrTyrArgThrGlyLysLeuLysAsnLeuArgAspGluProAsnMetLysTrp 300
Db 838 TGTCAAAATACCTATAGAGAGGACTAAATAAACTTCGAAAGCAACCTAATATAGATGAG 897
QY 301 SerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGln 320
Db 898 AGTATATTTAATGATTATCGAAGATATATGACTATTACTGTATAGATACTATCGCTCAA 957
QY 321 PheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyLeuValLysGly 340
|||||

Db 958 TTTTCTTTTATGATATAAGAGATACAAAGATTCATAAGCA-----AGAATAGGTGGC 1011
Qy 341 IleLysAsnGluLeuThrArgGluIleTyrThrGluIleAsnPheAspArgLeuPro 360
Db 1012 ATTAAACTGAACCTTACAGAGAAATTTATACAACTGAATAATTTTGACCGCTTACT 1071
Qy 361 GlnLeuArgValGlnProLeuLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPhe 380
Db 1072 TACCTTGAATTAACCAATTCGCTATATGGAATGGAATATAATTTAAACACGTTCCAGCGCT 1131
Qy 381 LysLeuPheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsn 400
Db 1132 AGATTTATTTTCAATTTTATGAGAACTTATATTTATACAAAATGAACGTCAGGGAAT 1191
Qy 401 ArgLeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
Db 1192 CGTTTGTAGTTGGAATTCGGAATCGTAATAGATCTACTTATCTACGACGGAATCGAAAT 1251
Qy 421 LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440
Db 1252 ATATATGGAGAAAGAACAGCTCCACCACCAACAAACITTTAATACATTTGAATCCTAT 1311
Qy 441 LysValSerIleValThrAspArgGlnSerProProValSerProIle----GlnProHis 459
Db 1312 AAAGTTTCAATTTGAATCTGATAGACAAAGTAACCTCTACTTCCCTTTTCCCTAAACATATAC 1371
Qy 460 PheIleLeuAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsnThrLeuLysTyr 479
Db 1372 TTTTCAATTAATCAAAATGAACTTTATTTAAATAATTCACCTAGTAATAAATTAACATAT 1431
Qy 480 SerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPheGlnPheProArgLys 499
Db 1432 TCAGCTGGGGGAATTTATCTAATGATAAAACCAACGATTTTCAATTTCTGTGATAA 1491
Qy 500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHis 519
Db 1492 AAAGACTGTAACCAATTAATTAATCAAAATGTTTACCAAGCTATAATAGTTATAGTCAT 1551
Qy 520 IleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIle 539
Db 1552 ATTTTATCCAGTTTCTTTTATTAATTAATTCCTATAAAATGGATAGCGCTAAATATA 1611
Qy 540 LeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSer 559
Db 1612 TTATATACAGGTGCATTAGATGACACACAGTAGTGTATTAAGATAATATGCAATATCA 1671
Qy 560 AspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLys 579
Db 1672 GATAAAATAATTAACAATGATCCAGCAATCAAAAGGTAACAGTCTTGATACAAACTCTAAG 1731
Qy 580 ValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArg 599
Db 1732 GTAATTTGAAGGACCTGGTCTACAGGAGGAAACTTGGTTTATTTACAAATGCAAGGCGT 1791
Qy 600 LeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyr 619
Db 1792 TTAGAGATTACATGTAGAACTCTTAATCTCACACAATCTTATACATTAAGCTTCGATAC 1851
Qy 620 AlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIle 639
Db 1852 GCTACAAATGGTCTGGAATACTCTTCCTTAATATATCTCTTACAAATACCAGGAGTAATA 1911
Qy 640 GlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGln 659
Db 1912 GGAATACCACTCAACGACTCAACACACTTTTTCGTGTACAAATTAATAATTTTACAA 1971
Qy 660 TyrGlyAspPheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsn 679
Db 1972 TAGCGAGATTTGGGTATTTCCAAATTTCCAAATGACATTAACATTTAAATCGAATAC 2031
Qy 680 IleProPheIlePheAsnArgAlaAspValSerAsnSerIleLeuIleAspLysIle 699
Db 2032 ATACCATTTATATTTAATCGTCAGATGTATCAAAATTTCAATTTTAAATCATTTGATAAAT 2091

Qy 700 GluPheIleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThr 719
Db 2092 GAATTTATACCAATTTACTTCTCTGTAGCCAAATAGAGAAAACAAATAATTAGAAACT 2151
Qy 720 IleGlnThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db 2152 ATCAAAACAAAATAAATACATTTTTCACAAATCATACAAA 2193
RESULT 3
CQ868309 2085 bp DNA linear PAT 13-SEP-2004
LOCUS Sequence 10 from Patent WO2004074462.
DEFINITION CQ868309
ACCESSION CQ868309.1 GI:51998355
VERSION CQ868309.1
KEYWORDS Bacillus thuringiensis
SOURCE Bacillus thuringiensis
ORGANISM Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE 1
AUTHORS Carozzi, N., Hargiss, T., Koziel, M.G., Duck, N.B. and Carr, B.
TITLE Delta-endotoxin genes and methods for their use
JOURNAL Patent: WO 2004074462-A 10 02-SEP-2004;
Athenix Corporation (US)
FEATURES
Location/Qualifiers
source 1..2085
/organism="Bacillus thuringiensis"
/mol_type="unassigned DNA"
/db_xref="taxon:1428"
CDS 1..2085
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAH33950.1"
/db_xref="GI:51998356"
/translations="MCGNTQYGDNFETPASADTAAVSAAGTIVSGTLLAGIGLTSI
SGPIGIIIGALIIISFTLITVFPAGQDKTWTQFKWGEIFVDTPLTESIKQLQT
LEGFROIQSNTALDDWRKLKRLQAPGLPSSALQQAALTAKRFENHNDVREIP
GFQLEWTKLLPIYAQAANPHNLLOQGAELADEWNADIHPSOIEPNAGTSDDYVKL
LKENIKYSNCANTYREGLNKLNEPNSISFENDYRRYRITVLTDTIAQFSYDIK
RKXDSIRIGGIKTELTREIYTEINFDRLTLEIQNLAIMETNLTRSGURLFSFL
ELIFYKNETYGNRLVGIANNRSTVATGTEIYBERTGPTTTLIPFESYKSVI
TDQVPTSPFPNIYFTINQILEYLNNSPNKLTYSAGNLSNDKKTDFQFPVKDC
KPIINCLPSYNSYSHLSQFSLEYSYKIGLALNLTLYGALCWTSSVVRNNAISD
KITMPLAKNSLIDTNSKYLEGPHGTGGLNVLQSOGRLEITCTCTNSTOSYVRLR
YATNAGNTLPNLSLTIPGVIGPPQRLNNTFGTINTNNLOYGDFGFPQFSTVTLPL
NRNPIFNRAVDVSNLIIIDKIEFIPITSSVRQNRKQKLETIQTINTPTNHTKN
TLNIEATNYDID"
ORIGIN
Alignment Scores:
Pred. No.: 3,51e-225 Length: 2085
Score: 3104.50 Matches: 598
Percent Similarity: 91.95% Conservative: 30
Best Local Similarity: 87.55% Mismatches: 52
Query Match: 80.24% Indels: 3
DB: 6 Gaps: 2
US-10-783-417-2 (1-735) x CQ868309 (1-2085)
Qy 52 MetCysGlnGlyAsnThrGlnTyrGlyAspAsnPheGluThrPheAlaSerAlaAspThr 71
Db 1 ATGTGTCAAGGAATACACAAATATGCTGATATATTTTCGAGACATTTGCTAGCTGATACA 60
Qy 72 IleAlaAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAlaGlyIleGlyGly 91
Db 61 ATGTCTCAGTTAGTCAGGACTATTGATATCCGCTACTCTGTTAGCCGTATAGCTGGG 120
Qy 92 LeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIleIleSerPheGlyThr 111
Db 121 CTCACCTTCTATATCCGACCGCATAGGAATAATAGTGTCTATAATAATATCTTTTGGTACC 180

Qy 112 LeuileThrValPheTyrProAlaGlyGluGlnAspLysThrValTyrThrGlnPheile 131
 Db 181 CTAATCACTGCTTTTGGCCCGCGGAGNACAAGACAAACAGATATGACAAATTTAT 240
 Qy 132 LysMetGlyGluIlePheValAspThrProLeuThrGluSerIleLysGlnLeuLysLeu 151
 Db 241 AAAATGGGAGAAATTTTGTGTGATACCGTTAAACAGAAACATAAACACAGCTAAAGTTA 300
 Qy 152 GlnThrLeuGluGlyPheArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyr 171
 Db 301 CAAACTTTAGAGGATTTAGACAATAATTAACAAGCTATATAACAGCATATAGATGATGG 360
 Qy 172 ArgLysLeuLysArgLeuGlnAlaProGlyLeuProSerSerAlaLeuGlnGlnAla 191
 Db 361 AGAAATTAATAAGACTACAGCTCCCTGGATTACCAACATCATCAGCATTAACAAGCT 420
 Qy 192 AlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPheIleArgGluIlePro 211
 Db 421 GCCTTGACTCTTAAATACGATTTGAGAACTGTCACAATGATTTTTCGAGAAATACCT 480
 Qy 212 GlyPheGlnLeuGluThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsn 231
 Db 481 GGTTCCAACTTGAACCTTATAAAGCTATTACTACCTATTATTATGCGCAAGCTGCTAAT 540
 Qy 232 PheHisLeuAsnLeuLeuGlnGlyAlaGluLeuAlaAspGluTyrAsnAlaAspIle 251
 Db 541 TTTCAATTAATTTATTACAAACAGGTGCTGAATTTGGCTGATGAATGGAATGCGAGATATA 600
 Qy 252 HisProSerGlnIleLeuProAsnAlaGlyThrSerAspTyrTyrLysLeuLeuLys 271
 Db 601 CATCTTCACAAATTTGAACCTTACTGCGAATCAGATGACTATTATAACTTTTAAAA 660
 Qy 272 GluAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsn 291
 Db 661 GAAATATACCTAAATATAGTAACCTATTGTGCAAAATACCTATAGAGAAGGACTAAATAA 720
 Qy 292 LeuArgAspGluProAsnMetLysTyrSerIlePheAsnAspTyrArgTyrMetThr 311
 Db 721 CTTGCAAAACCAACCTAATATAGATGAGATATATTAATGATTATCGAAGATATATGACT 780
 Qy 312 IleThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAsp 331
 Db 781 ATTACTGTATAGATATATGCTCAATTTCTTTTATGATATTAAGAGATACAAAGAT 840
 Qy 332 SerIleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThr 351
 Db 841 TCAATAGGA-----AGNATAGGTGGCATTAATACTGAACCTTACAGAGAAATTTATACA 894
 Qy 352 ThrGluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMet 371
 Db 895 ACTGAAATAAATTTTGACCGCTTACTTACCTTGAAATTTCAACCCAATCTCGCTATAATG 954
 Qy 372 GluTyrAsnLeuThrArgAlaSerPheLysLysPheSerPheLeuGluGlnPheIlePhe 391
 Db 955 GAATATATATTTAACACGTTTACGGGCTTAGATATTTTCATTTTATAGATGAACCTATATTT 1014
 Qy 392 TyrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIleSerAsnArgAspAlaPro 411
 Db 1015 TATACAAAATAGAACTGACGGGAATCGTTAGTTGGTATGCGAATCGTAATAGATCT 1074
 Qy 412 ThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySerProThrThr 431
 Db 1075 ACTTTATGCTACGACAGCACTGAAATATATATATATGAGAGAAACAGGTCACCCACAACA 1134
 Qy 432 LysThrIleArgProPheGluSerTyrLysValSerIleValThrAspArgGlnSerPro 451
 Db 1135 AAAACTTTAATACCATTTGATCTCTATAAAGCTTTCAATTTGTAACGTAGACAGTAAC 1194
 Qy 452 ProValSerProIle---GlnProHisPheIleLeuAsnGlnIleGluLeuTyrLeuAsn 470
 Db 1195 CCTACTTCCCTTTTCCCTAACATATATCTTCAATATATCAATGAACTTTATTTAAT 1254
 Qy 471 GlySerSerAsnAsnThrLeuLysTyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsn 490

Db 1255 AATTCACCTAGTAAATAAATTAACATATTACGCTGGGGGAATTTATCTAATGATAAAAA 1314
 Qy 491 ThrThrPheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCys 510
 Db 1315 ACAACTGATTTTCAATTTCTGTAAAAAAGACTGTAAACCAATTAATTAATCCAAATGT 1374
 Qy 511 SerProAsnPheAsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSer 530
 Db 1375 TTACCAAGCTATAATAGTATAGTATATTTATCCAGTTTCTTTATTAATTAATTC 1434
 Qy 531 TyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSer 550
 Db 1435 TATAAATTTGATTTAGCGCTAAATATATATATACAGGTGCATTTAGATGGACACAGT 1494
 Qy 551 SerValAspTyrGlyAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLys 570
 Db 1495 AGTGTTAATAGAAATAATGCAATATCAGATAAAATAATTAATCAATGATCCAGCAATCAA 1554
 Qy 571 GlyAsnAsnLeuAspThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsn 590
 Db 1555 GGTAAACAGCTCTTGATACAAACTCTAAGGTAAATGAAGGACCTGGTCATACAGAGGAAAC 1614
 Qy 591 LeuValTyrLeuGlnSerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThr 610
 Db 1615 TTGGTTTATTTACAAAGTCAAGGCGTTTAGAGATTACATGTAGAACTCTTAATTTCTACA 1674
 Qy 611 GlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsn 630
 Db 1675 CAATCTTATTACATTAGACTTCGATACCGTACAAATGGTGTGGAAATATCTTCTCTAAT 1734
 Qy 631 IleSerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsnAsnThrPhe 650
 Db 1735 ATATCTCTTACAAATACAGGAGTAAATAGGAATACCACCTCAACGACTCAACAAACACTTTT 1794
 Qy 651 SerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSer 670
 Db 1795 TCTGGTACAAATTAATAATTTACAAATACGAGATTTTGGGTATTTCCAAATTTCCAAAT 1854
 Qy 671 ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSer 690
 Db 1855 ACAGTAACTATACCTTTAAATCGAAACATACCAATTTATTAATCGTCAGATGTATCA 1914
 Qy 691 AsnSerIleLeuIleAspLysIleGluPheIleProIleThrSerSerMetHisGln 710
 Db 1915 AATTCAAATTTTAATCAATGATAAAATTTGAATTTATACCAATTTACTTCTCTGTACGCCAA 1974
 Qy 711 AsnArgGluLysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsn 730
 Db 1975 AATAGAGAAAAACAAAATTTAGAACTATCCAAACAAAATAAATACATTTTTCACAAAT 2034
 Qy 731 HisThrLys 733
 Db 2035 CATACAAA 2043
 RESULT 4
 BTTOXD1
 LOCUS 3543 bp DNA linear BCT 18-APR-2005
 DEFINITION Bacillus thuringiensis gene for 130 kDa delta-endotoxin.
 ACCESSION Y00423
 VERSION Y00423.1 GI:40351
 KEYWORDS delta-endotoxin; endotoxin.
 SOURCE Bacillus thuringiensis
 ORGANISM Bacillus thuringiensis
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
 REFERENCE 1 (bases 1 to 3543)
 AUTHORS Ward, E.S. and Ellar, D.J.
 TITLE Nucleotide sequence of a Bacillus thuringiensis var. israelensis gene encoding a 130 kDa delta-endotoxin
 JOURNAL Nucleic Acids Res. 15 (17), 7195 (1987)
 PUBMED 2821500
 REFERENCE 2 (bases 1 to 3543)

US-10-783-417-2 (1-735) x BTTOXD1 (1-3543)	Qy	1	MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleAspSerHisThrSerProTyr	20	1061.50	Length: 3543	Matches: 271	Conservative: 118	Mismatches: 278	Indels: 105	Gaps: 24
	Db	1	ATGAATCTTATCAAAATAAATAAATGAATATGAACATTAATGCTTCACAAAAAATAA	60							
	Qy	21	PhenProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro	40							
	Db	61	-----AATATATCTTAATAATTATACAGATATCCCAATAGAAAATAGTCCCAAAACAATTA	114							
	Qy	41	LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly	60							
	Db	115	TTACAAAGTACAAATTAATAAGATGGCTCAATATATGTGTACAGATATCAGCATATGGT	174							
	Qy	61	AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle	80							
	Db		----								
	Qy	1	MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleAspSerHisThrSerProTyr	20	1061.50	Length: 3543	Matches: 271	Conservative: 118	Mismatches: 278	Indels: 105	Gaps: 24
	Db	1	ATGAATCTTATCAAAATAAATAAATGAATATGAACATTAATGCTTCACAAAAAATAA	60							
	Qy	21	PhenProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro	40							
	Db	61	-----AATATATCTTAATAATTATACAGATATCCCAATAGAAAATAGTCCCAAAACAATTA	114							
	Qy	41	LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly	60							
	Db	115	TTACAAAGTACAAATTAATAAGATGGCTCAATATATGTGTACAGATATCAGCATATGGT	174							
	Qy	61	AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle	80							
	Db		----								
	Qy	1	MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleAspSerHisThrSerProTyr	20	1061.50	Length: 3543	Matches: 271	Conservative: 118	Mismatches: 278	Indels: 105	Gaps: 24
	Db	1	ATGAATCTTATCAAAATAAATAAATGAATATGAACATTAATGCTTCACAAAAAATAA	60							
	Qy	21	PhenProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro	40							
	Db	61	-----AATATATCTTAATAATTATACAGATATCCCAATAGAAAATAGTCCCAAAACAATTA	114							
	Qy	41	LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly	60							
	Db	115	TTACAAAGTACAAATTAATAAGATGGCTCAATATATGTGTACAGATATCAGCATATGGT	174							
	Qy	61	AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle	80							
	Db		----								
	Qy	1	MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleAspSerHisThrSerProTyr	20	1061.50	Length: 3543	Matches: 271	Conservative: 118	Mismatches: 278	Indels: 105	Gaps: 24
	Db	1	ATGAATCTTATCAAAATAAATAAATGAATATGAACATTAATGCTTCACAAAAAATAA	60							
	Qy	21	PhenProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro	40							
	Db	61	-----AATATATCTTAATAATTATACAGATATCCCAATAGAAAATAGTCCCAAAACAATTA	114							
	Qy	41	LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly	60							
	Db	115	TTACAAAGTACAAATTAATAAGATGGCTCAATATATGTGTACAGATATCAGCATATGGT	174							
	Qy	61	AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle	80							
	Db		----								
	Qy	1	MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleAspSerHisThrSerProTyr	20	1061.50	Length: 3543	Matches: 271	Conservative: 118	Mismatches: 278	Indels: 105	Gaps: 24
	Db	1	ATGAATCTTATCAAAATAAATAAATGAATATGAACATTAATGCTTCACAAAAAATAA	60							
	Qy	21	PhenProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro	40							
	Db	61	-----AATATATCTTAATAATTATACAGATATCCCAATAGAAAATAGTCCCAAAACAATTA	114							
	Qy	41	LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly	60							
	Db	115	TTACAAAGTACAAATTAATAAGATGGCTCAATATATGTGTACAGATATCAGCATATGGT	174							
	Qy	61	AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle	80							
	Db		----								
	Qy	1	MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleAspSerHisThrSerProTyr	20	1061.50	Length: 3543	Matches: 271	Conservative: 118	Mismatches: 278	Indels: 105	Gaps: 24

```
Qy 421 LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440
Db 1228 TTA-----AAATCTCTGGTTGGCAACAATATTATATT----- 1263
Qy 441 LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
Db 1264 ----- 1266
Qy 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
Db 1267 TTATTAAATGTCATAAGCTTAGATAATAAATATCTCTAAATGATTATAAATATTATTATAA 1326
Qy 477 -----LeuLysTyrSerAlaGlyGly 483
Db 1327 ATGGATTTTATAACTAATGTTACTAGACTTTTGGAGAAGAACTTACAGCAGGATCT 1386
Qy 484 SerLeuSerAsnTyrGln---AsnThrThrPheGlnPheProArgLysLysAspCys 502
Db 1387 GGGCAATAAATCTTATGATGATAATAAATAAATTTTCGGGTTACCAATCTTTAAACGAAGA 1446
Qy 503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
Db 1447 GAGAAATCAAGGAAACCCCTCTTTTCCAAACATATGATACTATAGTCATATTTTATCA 1506
Qy 523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuIleLeuAspThr 542
Db 1507 TTTATTAAGTCTTAGTATCCCTGGCAACATATAAACTCAAGTGTAT----- 1554
Qy 543 GlyValLeuGlyTyrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
Db 1555 ---ACGTTTCTGGACACACTCTAGTGTGATCTCTAAATAATACATTTATACACATTTA 1611
Qy 563 IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582
Db 1612 ACTACCCAAATTCAGCTGTAAGCGAAATTCACCTGGGACTGCTCTCAAGTTGTTCAA 1671
Qy 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle 602
Db 1672 GGACCTGGTCATACAGGAGGATTTAATT-----GATTTCAAGATCATTTTCAAAAT 1725
Qy 603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
Db 1726 AATGTCACACTCAAAATTTTCAACATCGTATTTTATAAGATTCGTTATGCTTCAAT 1785
Qy 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db 1786 GGAAGCCAAATCTCGAGCTGTATTAATCTTAGTATCCAGGGTAGCAGACTG--- 1842
Qy 643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
Db 1843 GGTATGGCACTCAACCCCACTTTTCTGTGTACAGATTATACGAATTTAAATAATAAAGAT 1902
Qy 663 PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
Db 1903 TTTCACTACTTGAATTTCTTAACAGAGTGAAATTTGCTCCAAATCAAAACATATCTCT 1962
Qy 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleAspLysIleGluPhe 701
Db 1963 GTGTTTAATCGTGGATGTATATACAAACACACAGTACTTATGTATAAATTTGAATTT 2022
Qy 702 IleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGln 721
Db 2023 CTGCCAATTACTCGTTCTATAAGAGAGGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2082
Qy 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db 2083 CAAATAATTAATACATTTTATGCAAAATCCTATAAAA 2118
RESULT 5
108083
LOCUS
DEFINITION Sequence 1 from Patent EP 0296870.
linear PAT 02-DEC-1994
```

```
ACCESSION I08083
VERSION I08083.1 GI:589204
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4451)
AUTHORS Ellar,D.J. and Ward,E.S.
TITLE New toxin-encoding DNA fragments from Bacillus thuringiensis
subsp. israelensis
JOURNAL Patent: EP 0296870-A1 1 28-DEC-1988;
FEATURES
    Location/Qualifiers
        1..4451
            /organism="unknown"
            /mol_type="unassigned DNA"
Alignment Scores:
    Pred. No.: 3,42e-70 Length: 4451
    Score: 1061.50 Matches: 271
    Percent Similarity: 50.3% Conservative: 118
    Best Local Similarity: 35.10% Mismatches: 278
    Query Match: 27.44% Indels: 105
    DB: 6 Gaps: 24
```

US-10-783-417-2 (1-735) x I08083 (1-4451)

```
Qy 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 891 ATGAATCTTATCAAAATAAATAAATGAATATGAACATTAATGCTTCACAAAAAATTA 950
Qy 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 951 -----AATATATCTCTAATAATATATAAGATATCAATAGAAAAATAGTCCAAAAACAATTA 1004
Qy 41 LeuGlnAsnThrAsnTyrLysGluThrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
Db 1005 TTCAAAAGTACAAATATATAAGATTGGCTCAATATGTGTCAACAGAAATCAGCAGTATGGT 1064
Qy 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIle 80
Db 1065 GGAGATTTTGAACATTTTATTGTAGT-----GGTGAACCTCAGTGCCCTATACTATT 1115
Qy 81 ValSerGlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGly 100
Db 1116 GTAGTGGGACCGTACTGACTGTTTCGGGTTCAACA-----CCCTTAGGA 1163
Qy 101 IleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120
Db 1164 CTT-----GCTTTAATAGTTTGTGTACATTAATACCAGTTCTTTTCCAGCCCAA 1214
Qy 121 GluGlnAspLysThrValThrThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db 1215 GACCAATTAACACA---TGGAGTGACTTTATAACACAAACTAAATAATATTAAAAAAA 1271
Qy 141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
Db 1272 GAAATAGCATCAACATATATAAGTAATCTAATAAATTTTAAACAGGTCGTTTAAATGTT 1331
Qy 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaPro 180
Db 1332 ATCAGCATTATCATTAATCACTTAAACATGGGAG-----AATAATCCA 1376
Qy 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGlu 200
Db 1377 AACCCAAATACTCAGGATGTAAAGGACACAAATCCAGTAGTAGTTTACCATTTCAA 1436
Qy 201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr 218
Db 1437 AATGTCATTCCAGAGCTGTAAACTCTGTCTCTCTAATCCTAGTAGTGGCGATTACTAT 1496
Qy 219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGln 238
```

Db 1497 AACATCTAGTATTCTCTAGTTATGACAAAGCAGCAAACTTACATCTGACTGATTAAAT 1556
Qy 239 GlnGlyValGluLeuAlaAspGluThrProAsnAlaAspIleHisProSerGlnIleGluPro 258
Db 1557 CAAGCGCTCAATTTTGAAGCGTATTTAAAAACAATCGCAATTCGATTATTAGAGCCT 1616
Qy 259 AsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSer 278
Db 1617 ---TTGCCAACACGCAATTCGATTATTATTCAGTATTGACTAAAGCTATAGAAGATTACACT 1673
Qy 279 AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro----- 296
Db 1674 AATTATTGTGTAACTATATAAAAGGATTAATTTAATTAAGACGCGCTGATAGT 1733
Qy 297 -----AsnMetLysTyrPheIlePheAsnAspTyrArgArgTyrMetThrIle 312
Db 1734 AATCTTGATGGAAATATAAATCGGAACACATACATACGTTATCGAACAATAATGACTACT 1793
Qy 313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer 332
Db 1794 GCTGTATTAGATGTTCTGCACTCTTCTCTAATATGATGAGTAATAT----- 1844
Qy 333 IleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThr 352
Db 1845 -----CCAAATAGGTGTCCTAATCTGAACTTACTCGAGAAATTTATCAGGTA 1889
Qy 353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
Db 1890 ---CTTAACCTCGAAGAAAGCCCTATAATATAT-----GACTTTCATATCAAGAG 1940
Qy 373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyr 392
Db 1941 GATTCACTACAGTAGACCG---CATTATTTACTTGGCTTGATCTCTTGAATTTTAT 1997
Qy 393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
Db 1998 GAAAAAGCGCAAACTACTCTCTAATAATTTTTCACCAGCCATTATAATATGTTTCATTAC 2057
Qy 402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
Db 2058 ACACCTTGATAATATATCCCAAAATCTAGTGTGTTTGGAAATCACAATGTAACGTGATAA 2117
Qy 421 LeuTyrGlyGluArgThrGlySerProThrTyrLysThrIleArgProPheGluSerTyr 440
Db 2118 TTA-----AAATCTCTGTTGGCAACAATATTATAT----- 2153
Qy 441 LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
Db 2154 ----- 2156
Qy 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
Db 2157 TTATTAATATGTCATAAGCTTAGATAATAATAATCTAAATGATTATAATATTAGTAAA 2216
Qy 477 -----LeuLysTyrSerAlaGlyGly 483
Db 2217 ATGGATTTTTTATAACTAATGTTACTAGACTTTTGGAGAAAGAACTTACAGCAGGATCT 2276
Qy 484 SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysLysAspCys 502
Db 2277 GGGCAATAACTATTATGATGTAATAAATAATATTTTCGGGTTACCAATCTTAAACGAAGA 2336
Qy 503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
Db 2337 GAGAATCAAGGAAACCCCTACCCCTTTTCCAAACATATGATAACTATAGTATATTATCA 2396
Qy 523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnGlnIleLeuAspThr 542
Db 2397 TTTATTAAAAGTCTGATATCCCTGCAACATATAAACTCAAGTGAT----- 2444
Qy 543 GlyValLeuGlyThrThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
Db 2445 ---ACGTTTCTGGACACACTCTAGTGTGATCCTAAAAATAACAATTTATACACATTTA 2501

Qy 563 IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582
Db 2502 ACTACCCAAATTCAGCTGTAAAGCGAATTCATCTGGAGCTGCTTCTTAAGGTTGTTCAA 2561
Qy 583 GlyProGlyHisThrGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle 602
Db 2562 GGAACCTGTGTCATACAGAGGGGATTTAAT-----GATTTCAAAGATCATTTCAAAT 2615
Qy 603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
Db 2616 ACATGTCACACTCAAAATTTTCAACAATCGTATTTTATAAGAAATCGTTATGCTTCAAT 2675
Qy 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyLeuPro 642
Db 2676 GGAAGCGCAAACTACTCGAGCTGTTTATAAATCTTAGTATCCAGGGGTAGCAGAACTG-- 2732
Qy 643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
Db 2733 GGTATGGCACTCAACCCCACTTTTCTGTGTACAGATTATACGAATTTAAATAATAAAGAT 2792
Qy 663 PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnAsnArgAsnIleProPhe 682
Db 2793 TTTCACTACATTAGAAATTTCTAACGAGGTGAATTTGCTCCAAATCAAAACATATCTCT 2852
Qy 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleIleAspLysIleGluPhe 701
Db 2853 GTGTTTAAATCGTTCGGATGTATATACAAACACACAGTACTTATGATAAATGGAAT 2912
Qy 702 IleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGln 721
Db 2913 CTGCCAATTACTCGTTCTATAAGAGAGATAGAGAGAAACAAAAATTAGAAACACGTACAA 2972
Qy 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db 2973 CAAATAATTAATACATTTTATGCAATTCCTATAAAA 3008
RESULT 6
I09103 LOCUS I09103 4451 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 1 from Patent WO 8810305.
ACCESSION I09103
VERSION I09103.1 GI:588188
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4451)
AUTHORS Ellar,D.J. and Ward,E.S.
JOURNAL Patent: WO 8810305-A 1 29-DEC-1988;
FEATURES
source 1..4451
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 3.42e-70 Length: 4451
Score: 1061.50 Matches: 271
Percent Similarity: 50.39% Conservative: 118
Best Local Similarity: 35.10% Mismatches: 278
Query Match: 27.44% Indels: 105
DB: 6 Gaps: 24
US-10-783-417-2 (1-735) x I09103 (1-4451)
Qy 1 MetAsnGlnAsnAsnAspAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 891 ATGAATCCTTATCAAAATAAATAAATGAATATGAACATTAATGCTTCCAAAAAATAATTA 950
Qy 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 951 -----AATATATCTAATAATATATACAAGATATCCAATAGAAAAATAGTCCAAAAACAATTA 1004

Qy 41 LeuGlnAsnThrAsnTyrLysGluTrpLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
Db 1005 TTAACAAGTACAAATTATAAAGATTGGCTCAATATGTGTCAACAGAATCAGCAGTATGGT 1064
Qy 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80
Db 1065 GGAGATTGTGAACCTTTATTGTAGT-----GGTGAACCTCAGTGCCTATACTATT 1115
Qy 81 ValSerGlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGly 100
Db 1116 GTAGTTGGGACCGTACTGACTGGTTTCGGGTTCAACA-----CCCTTAGGA 1163
Qy 101 IleIleGlyAlaIleIleIleSerPheGlyThrIleIleThrValPheTrpProAlaGly 120
Db 1164 CTT-----GCTTTAATAGGTTTGGTACATTAAATACAGTCTCTTTTCCAGCCCAA 1214
Qy 121 GluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db 1215 GACCAATCTAACACA---TGGAGTGACTTTTATAACACAACTAAATAATATTATAAAAAA 1271
Qy 141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
Db 1272 GAAATACATCAACATATATAAGTAAGTAACTAAATAAATTTTAAACAGTCTGTTAATGTT 1331
Qy 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaPro 180
Db 1332 ATCAGCAGCTTATCATATACCTTTAAACATGGGAG-----ATAAATCCA 1376
Qy 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaIleLeuThrLeuLysIleArgPheGlu 200
Db 1377 AACCCACAAATATCAGGATGTGAAGACACAAATCCAGCTAGTTCATTACCATTTTCAA 1436
Qy 201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr 218
Db 1437 AATGTCATTCAGAGCTGTAACTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1496
Qy 219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGln 238
Db 1497 AACATAGTATGTTATCTAGTATATGACAGCAGCAGCAAACTTACATCTGACTGTATTAAT 1556
Qy 239 GlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluPro 258
Db 1557 CAAGCCGTCAAATTTGAAGCGTATTTAAAAAACAATCGACAATTCGATTATTATAGAGCCT 1616
Qy 259 AsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSer 278
Db 1617 ---TTGCCAACAGCAATGTATTATCCAGTATTTGACTAAAGCTATAGAGATTACACT 1673
Qy 279 AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro----- 296
Db 1674 AATTATTGTGTACAACTTATAAAGAGATTAAATTTAATAAAGCAGCCCTGATAGT 1733
Qy 297 -----AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIle 312
Db 1734 AATCTTGATGGAATATAAATCTGGAACACATACATACGATATCGAACAATAAATGACTACT 1793
Qy 313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer 332
Db 1794 GCTGTATTAGATTGTGTGCACTCTTCTTAATATTATGATGTAGTAATAT----- 1844
Qy 333 IleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThr 352
Db 1845 -----CCAATAGGTGTCGAATCTGAACCTTACTCGAGAAATTTATCAGGTA 1889
Qy 353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
Db 1890 ---CTTAACCTCGAAGAACCCCTATATAATTTAT-----GACTTTCAATATCAAGAG 1940
Qy 373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnPheIlePheTyr 392
Db 1941 GATTCACTTACAGTAGACCG---CATTTATTACTTGGCTTGATTCTTTGAATTTTAT 1997

Qy 393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
Db 1998 GAAAAAGCGCAAACTACTCTTAATAATTTTTTCCACGCAATATAATATATGTTTCATTAC 2057
Qy 402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
Db 2058 ACATTGATATATATCCAAAAATCTAGTGTGTTGGAAATCACAATGTAATGATATAA 2117
Qy 421 LeuTyrGlyGluArgThrGlySerProThrTyrThrLysThrIleArgProPheGluSerTyr 440
Db 2118 TTA-----AAATCTCTGGTTTGGCAACAAATATTATATT----- 2153
Qy 441 LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
Db 2154 ----- 2156
Qy 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
Db 2157 TTATTAAATGTCAAGCTTAGATAATAAATATCTAAATGATTATAATAATATTAGTAAA 2216
Qy 477 -----LeuLysTyrSerAlaGlyGly 483
Db 2217 ATGGATTTTTTATAACTAATGGTACTAGACTTTTGGAGAGAACTTACACAGGATCT 2276
Qy 484 SerLeuSerAsnTyrGln---AsnThrThrPheGlnPheProArgLysLysAspCys 502
Db 2277 GGGCAAAATACTATGATGTAAATAAAAAATATTTTCGGGTTCACCAATTTCTTAAACGAAGA 2336
Qy 503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
Db 2337 GAGAAATCAAGAAACCCCTACCCCTTTTCCACATATGATAACTATAGTCATATTTATCA 2396
Qy 523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThr 542
Db 2397 TTTTATTAAGCTCTAGTATCTCCGACATATAAACTCAAGTGTAT----- 2444
Qy 543 GlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
Db 2445 ---ACGTTTGTGTCGACACACTCTAGTGTGATCTTAAATAAATACAAATTTATACACATTTA 2501
Qy 563 IleThrMetIleProAlaIleLysGlyAsnLeuAspThrAsnSerLysValIleGlu 582
Db 2502 ACTACCCAAATTCAGCTGTAAAGCGAAATTCACCTGGGACTGCTTCAAGGTTGTTCAA 2561
Qy 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle 602
Db 2562 GGACTGTGTATACAGGAGGGATTAAT-----GATTCAAGATCATTTCAAAT 2615
Qy 603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
Db 2616 ACATGTCAACACTCAAATTTTCAACATCGTATTATTATAAGAAATTCGTATGCTTCAAAT 2675
Qy 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db 2676 GGAAGCGCAATATCTCGAGCTGTATTATAATCTTAGTATCCAGGGGTAGCAGAACTG--- 2732
Qy 643 ProGlnArgLeuAsnAsnThrPheSerGlyTyrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
Db 2733 GGTATGGCACTCAACCCACCTTTCTGTYACAGATTATACGANTTTAAATATAAAGAT 2792
Qy 663 PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
Db 2793 TTTTCAGTACTTAGAATTTTCAACGAGGTGAAATTTGCTCCAAATCAAAACATATCTCTT 2852
Qy 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleAspLysIleGluPhe 701
Db 2853 GTGTTTAATCTCGATGTATATACAAACACACACAGTACTTATTGATAAAATTTGAATTT 2912
Qy 702 IleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGln 721
Db 2913 CTGCCAATTTACTCGTCTTATAGAGGATGAGAGAAACAAAAATTAGAACAGTACAA 2972
Qy 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733


```
Db 1347 -----CCAAATAGGTGTCCTCAATCTGAATCTTACTCGAGAAATTTATCAGGTA 1391
Qy 353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
Db 1392 ---CTTAACCTTCGAAGAAGCCCTATAAATATTAT-----GACTTTCAATATCAAGAG 1442
Qy 373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyr 392
Db 1443 GATTCACTTACAGTAGACCG---CAATTATTACTTGGCTTGATCTTTGCAATTTTAT 1499
Qy 393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
Db 1500 GAAAAAGCGCAACTACTCTCTAATAATTTTTCACCGCCATTATAATATGTTTCATTAC 1559
Qy 402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
Db 1560 ACATCTGATATATATATCCAAAAATCTAGTGTCTTTTGGAAATCACAATGTAACTGATATA 1619
Qy 421 LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440
Db 1620 TTA-----AAATCTCTGTTTGGCAACAATATTATAT----- 1655
Qy 441 LysValSerIleValThrAspArgGlnSerProValSerProIleGlnProHisPhe 460
Db 1656 ----- 1658
Qy 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
Db 1659 TTATTAAATGTCATAAGCTTAGATAATAAATATCTAAATGATTATAATAATATTAGTAAA 1718
Qy 477 -----LeuLysTyrSerAlaGlyGly 483
Db 1719 ATGGATTTTTTATACTAATAGTGTACTAGACTTTTGGAGAAAGAACTTACAGCAGGATCT 1778
Qy 484 SerLeuSerAsnTyrGln---AsnThrThrPheGlnPheProArgLysLysAspCys 502
Db 1779 GGGCAATAACTATGATGATAATAAATAATTTTCGGTACCACATTCCTTAACAGAGA 1838
Qy 503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
Db 1839 GAGAAATCAAGGAACCCCTACCCCTTTTCCAAATATGATAACTATAGTCATATTTATCA 1898
Qy 523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThr 542
Db 1899 TTTATTAAAAAGTCTTAGTATCCCTCGCAACATATATAAACTCAAGTGAT----- 1946
Qy 543 GlyValLeuGlyThrThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
Db 1947 ---ACGTTTCTGGACACACTCTAGTGTGTGATCTTAAANAATACAAATTTATACACATTTA 2003
Qy 563 IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582
Db 2004 ACTACCAAAATTCAGCTGTAAAGCGAATTCACCTGGGACTGCTTCTAAGTTGTCAA 2063
Qy 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle 602
Db 2064 GGACCTGGTCATACAGAGGGGATTTAATT-----GATTTCAAAGATCATTTCAAATTT 2117
Qy 603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
Db 2118 ACATGTCAACACTCAAAATTTTCAACAAATCGTATTTTATAAGAAATTCGTATGCTTCAAT 2177
Qy 623 GlyValGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db 2178 GGAAGCGCAAAATCTCGAGCTGTTATAAATCTTAGTATCCAGGGGTAGCGAACTG--- 2234
Qy 643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
Db 2235 GGTATGGCACTCAACCCCACTTTTCTGTTACAGATTATACGATTTAAATAATATAAGAT 2294
Qy 663 PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
Db 2295 TTTTCAGTACTTAGAATTTTCTCAACGAGGTGAAATTTGCTCCAAATCAAAACATATCTCTT 2354

Qy 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleIleAspLysIleGluPhe 701
Db 2355 GTGTTTAATCGTTCGGATGTATATACAAACACACAGTAGTACTTATTGATAAAATTGAATTT 2414
Qy 702 IleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGln 721
Db 2415 CTGCCAATTAATCTGTTCTATAGAGAGGATAGAGAGAAACAAATTTAGAACAGTACAA 2474
Qy 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db 2475 CAAATAATTAATACATTTTATGCAAAATCCTATAAAA 2510

RESULT 8
E01676
LOCUS E01676
DEFINITION DNA sequence encoding Bacillus thuringiensis insectivorous protein.
ACCESSION E01676
VERSION E01676.1 GI:2169929
KEYWORDS JP 1988230090-A/1.
SOURCE Bacillus thuringiensis
ORGANISM Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE 1 (bases 1 to 4934)
AUTHORS Komano, T. and Himeno, M.
INSECTICIDAL PROTEIN OF STRAIN BACILLUS THURINGIENSIS VAR
TITLE Patent: JP 1988230090-A 1 26-SEP-1988;
JOURNAL SUMITOMO CHEM CO LTD
COMMENT OS Bacillus thuringiensis
PN JP 1988230090-A/1
PD 26-SEP-1988
PF 19-MAR-1987 JP 1987066844
PI KOMANO TORU, HIMENO MICHIO
PC C12N15/00, A01N63/00, C12N1/20, C12P21/02, A01N63/02, C12N1/20,
CC C12R1:19;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=HD522;
FH Key Location/Qualifiers
FT 5'UTR 621..1368
FT CDS 1369..4911
FT /product="Bacillus thuringiensis insectivorous protein".
FT Location/Qualifiers
source 1..4934
/organism="Bacillus thuringiensis"
/mol_type="genomic DNA"
/db_xref="taxon:1428"

ORIGIN
Alignment Scores:
Pred. No.: 4,63e-70 Length: 4934
Score: 1060.50 Matches: 271
Percent Similarity: 50.39% Conservatives: 118
Best Local Similarity: 35.10% Mismatches: 278
Query Match: 27.41% Indels: 105
DB: 6 Gaps: 24

US-10-783-417-2 (1-735) x E01676 (1-4934)
Qy 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 1369 ATGAATCTTATCAAAATATAAAATGAATATGAATCAATTAATGCTTACAAAAAATA 1428
Qy 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 1429 -----AATATATCTTAATAATATTATCAAGATATCAATAGAAAATAGTCCAAACAAATTA 1482
Qy 41 LeuGlnAsnThrAsnTyrLysGluTrpLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
```


/note="HMMPfam hit to PF02899, Phage integrase, N-terminal

RBS	SAR-like domain"
gene	complement(5283). .5286)
	complement(5364). .5507)
	/gene="pBt006"
CDS	complement(5364). .5507)
	/gene="pBt006"
	/note="Similar to Bacillus anthracis pxiol-17 TR:Q9X2Y8 (EMBL:AF065404) (47 aa) fasta scores: E(): 2.1e-12, 68.0%
	id.in:47 aa

```

to in 4/ ca
/codon_start=1
/transl_table=11
/product="putative integral membrane protein"
/protein_id="CAD30068.1"
/db_xref="GI:21685415"
/db_xref="UniProt/TREMBL:Q8KNX2"
/translation="MAVLKDSNYEMTYNEIGFFKGIFWGLIFVVPFWSIMIATFIW
LCK"

```

misc feature

```
misc_feature
    LACK.
    complement(5370..5435)
    /gene="pBt006"
    /note="1 probable transmembrane helix predicted for pBt006
    by TMHMM2.0"
RBS
    complement(5515..5519)
    complement(6451..8160)
gene
    /gene="pBt007"
CDS
    complement(6451..8160)
    /gene="pBt007"
```

/note="Similar to *Bacillus anthracis* pxiol-16 TR:Q9X2Y7 (EMBL:AF065404) (569 aa) fasta scores: E(): 0, 96.13% id in 569 aa, and to *Bacillus thuringiensis* pxiol orf16-like protein TR:CA050562 (EMBL:AJ296638) (310 aa) fasta scores: E(): 6.5e-122, 99.67% id in 310 aa"

```

/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAD30069.1"
/db_xref="GI:21695416"

```

```

/db_xref="UniProt:TrEMBL:Q8KNX1"
/translation="MGSFPIIQNRSNGKTTNAAVVVTLKSWQKFAVAPEEKYSYVD
TYLEDILQHYERIADVEGVBEKAKPSFRRAMSTLAEKTRGHKLYGITNADLEV
FLYLHKCTKGVIPNTVTHMWEDYQYKEEFAYIQHSQFYIALKLNLIHNTISTEN
GLDGRYTKTLFWNEESTKANPYVSPVITKDFDKFLIAAKLTLFAMQOHTET
TLKSLDKQDSRGKNTFGGMYRLHKYKPHOIRAVIEELTALPCTGNPLFKICMKQ
KGVHTRKYTYLLSLHSDFLCSKEAGEEQRDPPTPKATYAKAKIEAVLQEMNIG
ELSDAMPTLHVLTCTHRQVRSVIRGMDMDRKEGYPTKIVITLKKLQHQTSQYQI
LDTAAKEGIVLAQHPKERNSDQAVENFLHYSMYSLHNIKFNKVRHALLKQK
FAPVPTESYVRYNTLKQEEITLFRKYAQDGVNLITALEIEMREKLKYVRGHKERTI
PSDVREWFIEALDKLPQEKLRVIELPKQFNLEFMRTFRLVRAGVITITAPDQVLHAM
PIK"

```

```

RBS
RBS
gene
CDS

EIK"
complement(8168. .8173)
8947. .8951
8959. .10653
/gene="pBt009"
8959. .10653
/gene="pBt009"
/note="Similar to Bacillus anthracis pxol-14 TR:Q9X2Y5
(EMBL:AF065404) (564 aa) fasta scores: E(): 1.6e-191.
```

ment Scores:			
. No.:	2.73e-68	Length:	127923
es:	1060.50	Matches:	271
ent Similarity:	50.39%	Conservative:	118
Local Similarity:	25.10%	Mismatches:	278
y Match:	27.41%	Indels:	105
		Gaps:	24

0-783-417-2 (1-735) x BTBPTOXIS (1-127923)

1 MetAsnGlnAsnAsnAsnAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
||||| :||| ||||| ||||| :||| :||| :||| :||| :||| :||| :|||
96528 ATGAATCCTTCATAAAATAAATAGTAATATGAAACATTAAAGCTTCACAAAAAATTA 9646

21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40

21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 10

Db	96468	: : : : : : : : : : : : : : :	-----AATATATCTATATTAATTATATCAAGATATTCATAGAAATATAGTCCAAACAATTA	96415
Qy	41	LeuGlnAsnThrAsnTyrIysGlyTrpLeuAenMetCysGlnGlyAsnThrGlnTyrGly	60	
Db	96414	: : : : : : : : : : : : : : : : : : : :	TTCAAAGTACAAATTTATAAAGATTGGCTCAATATGTGTCAACAAGATCAGCAGTATGGT	96355
Qy	61	AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle	80	
Db	96354	: : : : : : : : : : : : : : : : : : : :	GGAGATTTTGAAACTTTTATTGATAGT-----GGTGAACTCAGTGCCTATACTATT	96304
Qy	81	ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly	100	
Db	96303	: : : : : : : : : : : : : : : : : : : :	GTAGTTGGACCGTACTGACTGTTTCGGGTTTCAACA-----CCCTTTTGA	96256
Qy	101	IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly	120	
Db	96255	: : : : : : : : : : : : : : : : : : : :	-----GCTTTAATACGTTTGGTACATTAATACAGTCTCTTTTTCACGCCAA	96205

D6		96255	CIT-----GCTTAAAGGTTTGGTAGCATTAATACACAGTCTTTTTTCACGCCCA	96204
Qy	121	GlunInAspLysThrValTrpThrGlnPheLeuLysMetGlyGluIlePheValAspThr	140	
D6		96204	GACCAATCTAACACA---TGGAGTGACCTTATAACACAAAATCTAATAAAAAA	96148
Qy	141	ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle	160	
D6		96147	GAATAGCATCAACATATATAGAATAATGCCTAATAAAAAATTTAAACACGCTCGTTAATGTT	96088

Qy	161	LeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgueuGlnAlaPro	180
	:::	:::: :::	
Db	96087	ATCAGCACTTATCATTAATCACCTTAAACATGGGAG-----AATAATCCA	96043
Qy	181	GlyLeuProProSerSerAlaLeuGlnGlnAlaLaLeuThrLeuLysIleargphedu	200

96042 AACCCACAAATACTCAGGATCTAAGGACACAAATCCAGCTAGTTCATTACCAATTTTCAA 95983

Db

201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr 218

Qy

Db	95982	AATGTCATTCCAGAGCGCTTGTAACAACCTCTGTCCTCTTAATCCTAGTGAATTCGCGATTACTAT	95923
Qy	219	LysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGln	238
Db	95922	ACACATACCTAGTATATCTAGTTATATGCAACAGCAGCAAACTTCACTCGATCTGTTATTAAT	95863

Db	95922	AAACATACCTAGTATTATCTAGTTTATGCAACAGCAGCAAACTTATACCTCTGACTGATATAAA	95866
Db	95862	CAAGCCGTCATATTTGAAGCGTATTTAAAAAACAATTCGACATTTAGAGGCT	95803
Qy	239	gInGIyAlaGlULeUaAlaSpGLuTrPAsnAlaAaSpIleHisProSerGlnIleGLuPro	258

Qy	259	AenAlaGlyTyrSerAspPtyrTyrlsLeuLeuLysGluasnIleproLysTyrSer	278
Db	95802	---TTGCCAACGAATTGATTTATTCAGTATTGCCTAAGAAGCTATAGAAGATTACACT	95746

Qy 279 AaNTyTCyCysAlaasnThrTyTArGthrGlyLeuLysAsnLeuAraGspClupro----- 296
 95745 AaNTATTGTGTAAcAcTATTATAAAAAAGGATTAAATTTAATTTAAACACGACGCCTGATAGT 95686
 Db

Qy 297 -----AsnMetLysTrpSerIlePheAsnAspTyrArGArgTyrMetThrIle 312
|||::: |||::: |||::: |||::: |||:::
95685 AATCTTGATGGAAATATAAATCGGAACACATACATACGTTCGCAGAAAATGACTACT 95626
Db

Qy	313	ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer	332
Db	95625	GCTGTATTAGACTCTTTGTCACCTCTTTCTTAATTATGATGTAGGTAATAAT	95575

Qy 333 IleGlyGlyIleGluValIysGlyIleYAsnGluLeuThrArgGluIleYrThrThr 352
 95574 -----CCAAATAGGTGTCCAACTCGAACTTATCAGGTA 95530
 Db 353 GluIleAsmPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
 Qv

[illegible]

Db 95478 GATTCACTACAGTAGACCG---CATTATTACTTTGGCTTGATCTCTTGAATTTTAT 95422
Qy 393 -----ThrGluAsnThrAsnPhGlyAsnArg----- 401
Db 95421 GAAAAAGCGCAAACTACTCTCTAATATTTTTCACCGCAATTAATAATATGTTTCATTAC 95362
Qy 402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
Db 95361 ACACCTGATATATATCCCAAAATCTAGTGTGTTTGGAAATCAACATGACATGATGATAA 95302
Qy 421 LeuTyrGlyGluArgThrGlySerProThrThrIleThrIleThrIleThrIleThrIle 440
Db 95301 TTA-----AAATCTCTGTGTTGGCAACAAATATTTATATT----- 95266
Qy 441 LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
Db 95265 ----- 95263
Qy 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
Db 95262 TTATTAATGCTCAAGCTTAGATAATAAATTAATTAATGATTAATAATATAGTATAA 95203
Qy 477 -----LeuLysTyrSerAlaGlyGly 483
Db 95202 ATCGATTTTTTATACTAATAGTACTAGACTTTTGGAGAAAGCACTACAGCAGATCT 95143
Qy 484 SerLeuSerAsnTyrGln---AsnThrThrPheGlnPheProArgLysLysAspCys 502
Db 95142 GGGCAAAATACTTATGATGATAATAAATAATTTTCGGGTACCAATCTTAAACGAAGA 95083
Qy 503 AsnLeuValIleAspProGlyCysSerProAsnPhenAsnTyrSerHisIleLeuSer 522
Db 95082 GAGAATCAAGGAACCCCTACCTTTTCCCAACATATGATACTATAGTATATTTTATCA 95023
Qy 523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnIleLeuAspThr 542
Db 95022 TTTTATAAAGCTTAGTATCCCTGCAACATATAAACTCAAGTGAT----- 94975
Qy 543 GlyValLeuGlyTrpThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
Db 94974 ---ACGTTTCTGGACACACTCTAGTGTGATCTTAAATAACAATTTATACACATTTA 94918
Qy 563 IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582
Db 94917 ACTACCCAAATTCAGCTGTAAAGCGAAATTCATCTGGGACTGCTTCAAGGTGTTCAA 94858
Qy 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle 602
Db 94857 GGACCTGGTCTACAGAGGGGATTTAAAT-----GATTCAAAGATCATTTCCAAAT 94804
Qy 603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
Db 94803 ACATGTCAACACTCAAAATTTTCAACATCGATTTATATAGAAATCGTTATGCTTCAAT 94744
Qy 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db 94743 GGAAGCGCAATACTCGAGCTGTATATAATCTTAGTATCCAGGGGTAGCAGAACTG--- 94687
Qy 643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
Db 94686 GGTATGGCACTCAACCCCACTTTTCTCTGTGTACAGATTATACGAATTTAAATAATAAAGAT 94627
Qy 663 PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
Db 94626 TTTCAGTACTTGAATTTCTTAACGAGGTGAATTTGCTCTCAAAATCAAAACATATCTCTT 94567
Qy 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleAspLysIleGluPhe 701
Db 94566 GTGTTTAATCGTTCGGATGTATATACAAACACACAGACTACTTATGATAAAATGAAAT 94507
Qy 702 IleProIleThrSerSerMetHisGlnAsnArgGluLysGlnIleLeuGluThrIleGln 721
Db 94506 CTGCCAATTAATCTCGTTCTATAAGAGGAGATAGAGAGAAACAAAATTTAGAAACAGTACAA 94447

Qy 722 ThrIysIleAsnThrPhePheThrAsnHisThrIys 733
Db 94446 CAAATAATTAATACATTTTATGCAAAATCCTATAAA 94411
RESULT 10
BACMSQB
LOCUS BACMSQB 3753 bp DNA linear BCT 26-APR-1993
DEFINITION B.Churingiensis insecticidal endotoxin gene, complete cd.
ACCESSION M12662
VERSION M12662.1 GI:143228
KEYWORDS
SOURCE
ORGANISM
Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE
AUTHORS 1 (bases 1 to 3753)
Thorne, L., Garduno, F., Thompson, T., Decker, D., Zounes, M., Wild, M., Walfield, A.M. and Pollock, T.J.
TITLE Structural similarity between the lepidoptera- and diptera-specific insecticidal endotoxin genes of Bacillus thuringiensis subsp. 'kurstaki' and 'israelensis'
J. Bacteriol. 166 (3), 801-811 (1986)
JOURNAL 3011746
PUBMED
COMMENT Original source text: B.thuringiensis subspecies israelensis DNA.
FEATURES
Location/Qualifiers
1..3753
/organism="Bacillus thuringiensis"
/mol_type="genomic DNA"
/db_xref="taxon:1428"
941..2968
/note="insecticidal endotoxin (put.); putative"
/codon_start=1
/transl_table=11
/protein_id="AAA22614.1"
/db_xref="GI:143228"
/translation="MNPQNKNVEYIFNAPNGFSKNYSRYPLANKPNQPLKNTNY KDLNVQDQNNQGNAGNFASSETIVGVSGAIIIVGTMGFAFAAALAAAGISFOTL LPIFWQSGDIPANVQDLNIGRPIQIEIDKNI INVLTIVTP IKQLDKYQFFPKWE PARTHAKA VHDLTFTLEPTIDKDLMLKNASVRIPTLPAYAQIATWHLNLLKHAATYNIWLQNGQINPSTFNSNNYGYLARKIQEYTDYCIOTYNAGLTMLRTNTNATW MYNRYLEMTLVLDLAIAPNDPEKYPGVKSELIREVTNVNSDFRTITELENG LTRNPTLTFFPNRGRFYTRNSRILDPYDIFSTGNQMAFTNDRDNLINGAVHNGI ISQDTSVFPYNNKRPIDKVEIRHREYSDIIYEMIFPSNSEVSEFNSNSTIENNYK RITQIMLPKOTWKNEEYCHTLYSKTNDYIFSVRRERVRVAFSWHTS VDFONTIDLD NTIOIHALKALKVSSDSKI VKGPGHTGGDLVLKDSMDRFRVFLKNSVQYQVRIYA TNAPKTVFELGTIDTISVELSPSTSRQNPNDLTLYADEGYVTFPTVPNKTFEGEDT LMLTYGTPNHSNIYIDKIEFIPITQSVLDYTEKQIEKTOKI VNDLFVN"
ALIGNMENT Scores:
Pred. No.: 7,08e-58 Length: 3753
Score: 897.50 Matches: 247
Percent Similarity: 48.32% Conservative: 126
Best Local Similarity: 31.99% Mismatches: 270
Query Match: 23.20% Indels: 129
DB: 1 Gaps: 30
US-10-783-417-2 (1-735) x BACMSQB (1-3753)
Qy 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 941 ATGAATCCATATCAAAATAAGAAATGAATATGAAATATTCATGCTCCATCAATGTTT 1000
Qy 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 1001 -----AGCAAGTCTAATAACTATTCTAGATATCCATAGCAATAAGCAATCAACCA 1054
Qy 41 LeuGlnAsnThrAsnTyrLysGluTrpLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
Db 1055 CTGAAAACACGAATTACAAAGATTGGCTCAATGTGTGTCAGAGATAATCAACAATATGC 1114
Qy 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80

[illegible][illegible]


```

DEFINITION  Bacillus thuringiensis subsp. israelensis plasmid pBtoxis.
ACCESSION   AL731825
VERSION     AL731825.1 GI:21685410
KEYWORDS
SOURCE      Bacillus thuringiensis serovar israelensis
ORGANISM    Bacillus thuringiensis serovar israelensis
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
            cereus group.
REFERENCE   1 (bases 1 to 127923)
AUTHORS    Barry,C., O'Neil,S., Ben-Dov,E., Jones,A.F., Murphy,L., Quail,M.A.,
            Holden,M.T., Harris,D., Zaritsky,A. and Parkhill,J.
TITLE       Complete sequence and organization of pBtoxis, the toxin-coding
            plasmid of Bacillus thuringiensis subsp. israelensis
JOURNAL     Appl. Environ. Microbiol. 68 (10), 5082-5095 (2002)
PUBMED     12324359
REFERENCE   2 (bases 1 to 127923)
AUTHORS    Parkhill,J.
TITLE       Direct Submission
SUBMITTED   (19-APR-2002) Submitted on behalf of the pBtoxis
            sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
            Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
COMMENT     Notes:
            Details of pBtoxis sequencing at the Sanger Centre are available on
            the World Wide Web.
            (URL, http://www.sanger.ac.uk/Projects/B\_b\_thuringiensis/).
FEATURES             Location/Qualifiers
     source            1..127923
                        /organism="Bacillus thuringiensis serovar israelensis"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:1430"
                        /note="plasmid pBtoxis"
     RBS               1145..1149
     gene              1163..2089
                        /gene="pBt001"
     CDS               1163..2089
                        /gene="pBt001"
                        /note="Similar in part to Bacillus anthracis pxol-49
                        TR:Q9X319 (EMBL:AF065404) (227 aa) fasta scores: E():
                        8.9e-44, 78.48% id in 158 aa"
                        /codon_start=1
                        /transl_table=11
                        /protein_id="CAD30064.1"
                        /db_xref="GI:21685411"
                        /db_xref="UniProt/TREMBL:Q8KNX5"
                        /translation="MSTHLFSQNFQKVGKFFDGLFIKLENERKELPSSALATY
                        IMLHFCNDIGLPRFQIMDLAKSGIPYTTITGQVCLERKLVREIPVGNRTVYE
                        IVDVALYNHTAETDMINISLVSFIPFYLQTTILSLVKARDNGIMLLELINT
                        FSKIGMEHYKHIEDPEITRKMAFLKELNRNAKVRQYIEIVKPIKFNVDLKKK
                        SSSRTVRVRQVQVIEIKFNVIFNSNCVIENDKKELHRIPIAKRKEAVSRLLKHMGG
                        ALKKDKENIMTAFROEIVDITIELYLPKQKN"
     repeat_unit       2246..3106
                        /note="IS240"
                        /complement(2246..2262)
     repeat_unit       /rpt_type=INVERTED
     gene              2337..3044
                        /gene="pBt003"
     CDS               2337..3044
                        /gene="pBt003"
                        /note="Similar to Bacillus thuringiensis insertion element
                        IS240-a protein TR:Q45766 (EMBL:M23740) (235 aa) fasta
                        scores: E(): 2.2e-92, 99.57% id in 235 aa, and to
                        Mycobacterium fortuitum, transposase tnp tnp6100 TR:Q49185
                        (EMBL:X53635) (254 aa) fasta scores: E(): 1.4e-37, 48.05%
                        id in 231 aa"
                        /codon_start=1
                        /transl_table=11
                        /product="insertion element IS240 protein"
                        /protein_id="CAD30065.1"
                        /db_xref="GI:21685412"
                        /db_xref="GOA:Q8KNX4"
                        /db_xref="InterPro:IPR001584"
                        /db_xref="UniProt/TREMBL:Q8KNX4"
                        /translation="MEKENIFKWKHYQADMILTVRWYRLYNLSFRDLVEMMEERGSL
                        LSHTTIMRWVHQYGPPELNERIRKHLKRTNDSWRVDETYIKIGENMYLYRAVDSEGNT
                        LDFYLSKKRDAKAAKFLKALASPHVTKPRVITVDGKNKAYPAVIRELNKESIPYGM
                        PLRVKYNLNMMIEQDHRFIKKIRNMLGLKSNQTAVKTAGIEMHVMVKGQQLKLRQAQ
                        3813..4208
                        /gene="pBt004"
                        /note="HMPFfam hit to PF00665, Integrase core domain"
                        4365..4381
                        /rpt_type=INVERTED
                        /complement(4634..5275)
                        /gene="pBt005"
                        /complement(4634..5275)
                        /gene="pBt005"
                        /note="Similar to N-terminus of Bacillus anthracis pxol-18
                        TR:Q9X2Y9 (EMBL:AF065404) (315 aa) fasta scores: E():
                        4.2e-56, 84.15% id in 183 aa, and to Bacillus halodurans
                        Bh2364 protein TR:Q9KACS (EMBL:AP001515) (378 aa) fasta
                        scores: E(): 1.6e-18, 35.45% id in 189 aa, and weakly to
                        Lactobacillus delbrueckii integrase/recombinase orf2
                        TR:Q48538 (EMBL:Z50864) (333 aa) fasta scores: E(): 6.3,
                        28.88% id in 90 aa, and to Bacillus thuringiensis
                        resolvase tnpI SW:TNRI_BACTU (P10020) (284 aa) fasta
                        scores: E(): 8.5, 23.88% id in 180 aa"
                        /codon_start=1
                        /transl_table=11
                        /product="integrase/recombinase family protein"
                        /protein_id="CAD30067.1"
                        /db_xref="GI:21685414"
                        /db_xref="GOA:Q8KNX3"
                        /db_xref="InterPro:IPR004107"
                        /db_xref="UniProt/TREMBL:Q8KNX3"
                        /translation="MSKNINLEETLAFAFSAVLIIEGRKQSTIKRYAYDIKDFYKWLRA
                        NKGLHHSKSELSEADYQTFSELEDKRRYSQKTRHRIWVVLKHLMFELCVISPDGG
                        INLSLIPQSLNDNDFTIMEEKLKQTVLSTKGLTERQAKYRPLTMDKNCIINLVV
                        NYGLSLVSLNNSHINFARNIWCYVKKIDTLNRDILLFKOLEIVAFHRTSP"
                        /complement(4985..5251)
                        /gene="pBt005"
                        /note="HMPFfam hit to PF02899, Phage integrase, N-terminal
                        SAM-like domain"
                        /complement(5283..5286)
                        /complement(5364..5507)

```

```

LSHTTIMRWVHQYGPPELNERIRKHLKRTNDSWRVDETYIKIGENMYLYRAVDSEGNT
LDFYLSKKRDAKAAKFLKALASPHVTKPRVITVDGKNKAYPAVIRELNKESIPYGM
PLRVKYNLNMMIEQDHRFIKKIRNMLGLKSNQTAVKTAGIEMHVMVKGQQLKLRQAQ
2538..2933
/gene="pBt003"
/note="HMPFfam hit to PF00665, Integrase core domain"
3090..3106
/rpt_type=INVERTED
3521..4381
/note="IS240"
complement(3521..3535)
/rpt_type=INVERTED
3612..4319
/gene="pBt004"
3612..4319
/gene="pBt004"
/note="Similar to Bacillus thuringiensis insertion element
IS240-a protein TR:Q45766 (EMBL:M23740) (235 aa) fasta
scores: E(): 3.5e-91, 99.14% id in 235 aa, and to
Mycobacterium fortuitum, transposase tnpA or tnp6100
TR:Q49185 (EMBL:X53635) (254 aa) fasta scores: E():
1.1e-37, 48.05% id in 231 aa"
/codon_start=1
/transl_table=11
/product="insertion element IS240 protein"
/protein_id="CAD30066.1"
/db_xref="GI:21685413"
/db_xref="GOA:Q8KH55"
/db_xref="InterPro:IPR001584"
/db_xref="UniProt/TREMBL:Q8KH55"
/translation="MEKENIFKWKHYQADMILTVRWYRLYNLSFRDLVEMMEERGSL
LSHTTIMRWVHQYGPPELNERIRKHLKRTNDSWRVDETYIKIGENMYLYRAVDSEGNT
LDFYLSKKRDAKAAKFLKALASPHVTKPRVITVDGKNKAYPAVIRELNKESIPYGM
PLRVKYNLNMMIEQDHRFIKKIRNMLGLKSNQTAVKTAGIEMHVMVKGQQLKLRQAQ
3813..4208
/gene="pBt004"
/note="HMPFfam hit to PF00665, Integrase core domain"
4365..4381
/rpt_type=INVERTED
/complement(4634..5275)
/gene="pBt005"
/complement(4634..5275)
/gene="pBt005"
/note="Similar to N-terminus of Bacillus anthracis pxol-18
TR:Q9X2Y9 (EMBL:AF065404) (315 aa) fasta scores: E():
4.2e-56, 84.15% id in 183 aa, and to Bacillus halodurans
Bh2364 protein TR:Q9KACS (EMBL:AP001515) (378 aa) fasta
scores: E(): 1.6e-18, 35.45% id in 189 aa, and weakly to
Lactobacillus delbrueckii integrase/recombinase orf2
TR:Q48538 (EMBL:Z50864) (333 aa) fasta scores: E(): 6.3,
28.88% id in 90 aa, and to Bacillus thuringiensis
resolvase tnpI SW:TNRI_BACTU (P10020) (284 aa) fasta
scores: E(): 8.5, 23.88% id in 180 aa"
/codon_start=1
/transl_table=11
/product="integrase/recombinase family protein"
/protein_id="CAD30067.1"
/db_xref="GI:21685414"
/db_xref="GOA:Q8KNX3"
/db_xref="InterPro:IPR004107"
/db_xref="UniProt/TREMBL:Q8KNX3"
/translation="MSKNINLEETLAFAFSAVLIIEGRKQSTIKRYAYDIKDFYKWLRA
NKGLHHSKSELSEADYQTFSELEDKRRYSQKTRHRIWVVLKHLMFELCVISPDGG
INLSLIPQSLNDNDFTIMEEKLKQTVLSTKGLTERQAKYRPLTMDKNCIINLVV
NYGLSLVSLNNSHINFARNIWCYVKKIDTLNRDILLFKOLEIVAFHRTSP"
/complement(4985..5251)
/gene="pBt005"
/note="HMPFfam hit to PF02899, Phage integrase, N-terminal
SAM-like domain"
/complement(5283..5286)
/complement(5364..5507)

```

```

misc_feature
repeat_unit
repeat_unit
repeat_unit
gene
CDS
misc_feature
repeat_unit
gene
CDS
misc_feature
RBS
gene

```

```

CDS
/genes="pbt006"
/complement (5364. .5507)
/genes="pbt006"
/notes="Similar to Bacillus anthracis pxol-17 TR:Q9X2Y8
(EMBL:AF065404) (47 aa) fasta scores: E(): 2.1e-12, 68.08%
id in 47 aa"
/codon_start=1
/transl_table=11
/product="putative integral membrane protein"
/protein_id="CAD30068.1"
/db_xref="GI:21685415"
/db_xref="UniProt/TrEMBL:Q8KNX2"
/translations="MAVLKDSNIEMTYNEIGFFKGFWMGLFVVPFWSIMIAITFIW
LCK"
misc_feature
complement (5370. .5435)
/genes="pbt006"
/notes="1 probable transmembrane helix predicted for pbt006
by TMHMM2.0"
RBS
complement (5515. .5519)
complement (6451. .8160)
/genes="pbt007"
complement (6451. .8160)
/genes="pbt007"
/notes="Similar to Bacillus anthracis pxol-16 TR:Q9X2Y7
(EMBL:AF065404) (569 aa) fasta scores: E(): 0, 96.13% id
in 569 aa, and to Bacillus thuringiensis pxol1 orf16-like
protein TR:CAC50562 (EMBL:AJ296638) (310 aa) fasta scores:
E(): 6.5e-122, 99.67% id in 310 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAD30069.1"
/db_xref="GI:21685416"
/db_xref="UniProt/TrEMBL:Q8KNX1"
/translations="MKSFLIQNSLNGKITNAVVLKSWQFVAAPBEKYSYKD
TYLEDILQMTVEERIDAVGVFEKAKPSFRAMSTLAETGRKHKLYGITNADLEV
FYLLHKKQTYEIPNVITIMMEDYKQYKEEFAYIOHSQFYIALKKLNHNTISIE
GLDGRYTIKLTLMNETSEKAMPYVVISPVITKDFKLSIAAKKLFDIAMOQHTET
TLKRSUDKODERNKTHFGGMVRFHKKYHQIRAVIEULTTALPCTGNPLFKICKMQ
KGVHTKRYTTLTSLHSPLCSKEAGEQHROPFPKATYAKKAEIEAVLQBMNIG
ELSDMNKIHLVLTICHRQIRSVIRGLRDVDRKGGYPTKIYTTLLKLLHQTSQVQI
LDTAKEGILYLAQHPKERNSDREQVNFGLHYSMSYSLHNIKMFNRVHALLKQK
FAPVTEESYNYLYQOETLFRKYADQGVNLHAYIALEIEMREKLKVRGHKERTI
PSDVRWFTEAIDKLPQEXLRVIELPQFNLLSFMRTERFLVRAGVITITAPDQVLHAM
EIK"
complement (8168. .8173)
RBS
8947. .8951
RBS
8959. .10653
Gene
/genes="pbt009"
8959. .10653
/genes="pbt009"
/notes="Similar to Bacillus anthracis pxol-14 TR:Q9X2Y5
(EMBL:AF065404) (564 aa) fasta scores: E(): 1.6e-191,
Alignment Scores:
Pred. No.: 5 88e-56 Length: 127923
Score: 897.50 Matches: 247
Percent Similarity: 48.32% Conservative: 126
Best Local Similarity: 31.99% Mismatches: 270
Query Match: 23.20% Indels: 129
DB: 1 Gaps: 30

US-10-783-417-2 (1-735) x BTPBTOXIS (1-127923)
QY 1 MetAsnGlnAsnAsnAsnAsnGluTyrGluIleAspSerHisThrSerProTyr 20
Db 39061 ATGAATCCATATCAAAATAAGATGAATATGAAATATTCATGCTCCATCCAAATGGTTT 39120
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 39121 -----AGCAAGTCTAATACTATTCTAGATATCCATTAGCAAAATACGCCAACCA 39174
QY 41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60

```

Db 40117 AGAAATTCGAGACATCTTGTATGATATTTTCTTTTACAGTAACAGATG 40176
QY 403 ValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyr 422
Db 40177 -----GCCTTACATACATACTAATGATGATGCGCAACATAATCTGG 40215
QY 423 GlyGluArgThrGly-----SerProThrThrIleArgProPheGluSer 439
Db 40216 GGAGCGGTTTCAGAAATATATTTCTCAAGACATCCAAAGATATTCCTTTTATAGA 40275
QY 440 Tyr-----LysValSerIleValThrAspArgGlnSerProProValSerPro 455
Db 40276 AACAAACCTATTGATAAGTCGAAATTTGTACAGATGAGAGTACTCAGATATAATA--- 40332
QY 456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn 475
Db 40333 -----TATGAAATGATATTTTTCGAATAGCAGTAA 40365
QY 476 ThrLeuLysTyrSerAlaGlyGlySerLeu---SerAsnTyrGlnAsnThrThrPhePhe 494
Db 40366 GTATTTTCGATATTCATCCATTCACATAGAAATATATATAAAGAACTGATCTTAT 40425
QY 495 GlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPhe 514
Db 40426 ATGATTCCTCAAAACAAACATCGMAAAAT----- 40452
QY 515 AsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGly 534
Db 40453 AAAGAAATATGTCATCTATCTATGATATAAAACATGATAATATATATTTTCAGTAGTT 40512
QY 535 LeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArg 554
Db 40513 AGAAGAAGAGAGAGTT-----GCATTAGTTGGACACATATAGTGTTGATTTC 40563
QY 555 TyrAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeu 574
Db 40564 CAAATACATAGATTAGATAACATCACCACCAATCCAGCTCTAAAGCTTTGAAGTA 40623
QY 575 AspThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeu 594
Db 40624 AGTTCTGATTCGAAATTTGTGAAAGTCTCTGTCACACAGGTGGAGACTTGGTAATCTT 40683
QY 595 GlnSerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPhe 614
Db 40684 AAGATAGTAGTATGATTTAGAGTTAGATT---TTAAATAATGTTTCTCGACATATCAA 40740
QY 615 IleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThr 634
Db 40741 GTACGATTCGTTATGCTACTAATGCT-----CCAAAGACAAACAGTATTC 40785
QY 635 IleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsn 654
Db 40786 TTAACCGGAATAGATACATAAGTGTGAG---CTCCCTAGTACCCTTCCCGCCAAAC 40842
QY 655 TyrAsn-----AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThrVal 672
Db 40843 CCAATGCTACAGATTAAACATATGCAATTTTGGATATGTAACATTTTCCAAGAACAGTT 40902
QY 673 -----ThrLeuProLeuAsnArgAsnIleProPheIle 683
Db 40903 CCAATAAACAATTTGAAGGAGAGACACATTTATTAATGACC-----TTA 40947
QY 684 PheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIleGluPheIlePro 703
Db 40948 TATGGTACACCAATCATTCATATATAATA---TATATTCAGCAAAATCGAATTTATTTCCA 41004
QY 704 IleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuThrIleGlnThrLys 723
Db 41005 ATCACTCAATCTCTATTAGATTATACAGAGAAGCAAAATATAGAAAAACACAGAAAAA 41064
QY 724 IleAsnThrPhePheThrAsnHisThrLysThrLeu 735
Db 41065 GTGAATGATTTATTTGTTAATTAACAAAGATTCTT 41100

RESULT 12

E00614
LOCUS
DEFINITION
E00614 DNA encoding a polypeptide having insecticidal activity(BTI
endotoxin).

ACCESSION

E00614 GI:2168893
VERSION
E00614.1 GI:2168893
KEYWORDS
JP 1986005098-A/1.
SOURCE
Bacillus thuringiensis
ORGANISM
Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
cereus group.

REFERENCE

1 (bases 1 to 3756)
AUTHORS
Aran,E.U. and Toomasu,J.P.
TITLE
INSECTICIDAL POLYPEPTIDE
JOURNAL
Patent: JP 1986005098-A 1 10-JAN-1986;
SHINTOKO CORP

COMMENT

OS Bacillus thuringiensis
PN JP 1986005098-A/1
PD 10-JAN-1986
PF 21-FEB-1985 JP 1985031647
PR 22-FEB-1984 US 84 582506, 22-JAN-1985 US 85 693556 PI
ARAN EMU UORUFUIRUDO, TOOMASU JIRI POROTSUKU PC
C07K15/04,A01N63/02,C07H21/04,C12N15/00//C12P21/00, C12N15/00, PC
C12R1:19);
CC strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=Israelensis ONR-60A;
CC *source: clone=BTIA4-1;
FH Key Location/Qualifiers

FEATURES

source
1..3756
/organism="Bacillus thuringiensis"
/mol_type="genomic DNA"
/db_xref="taxon:1428"
Alignment Scores:
Pred. No.: 1.49e-56 Length: 3756
Score: 880.00 Matches: 245
Percent Similarity: 47.85% Conservative: 123
Best Local Similarity: 31.86% Mismatches: 279
Query Match: 22.74% Indels: 122
DB: 6 Gaps: 31
US-10-783-417-2 (1-735) x E00614 (1-3756)

ORIGIN

1 MetAsnGlnAsnAsnAspAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
QY
Db 941 ATGAATCCATATCAAAATAGAAATAGATATGAAATATTCATGCTCCATCGATGTTT 1000
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 1001 -----AGCAAGTCTAATACTATTTCTAGATATCCATTAGCAAAATAGCAATCAACCA 1054
QY 41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
Db 1055 CTGAAAAACACCAATTAACAAAGATTGGCTCAATGTGTGTCAAGATAATCAACAATATGGC 1114
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80
Db 1115 AATAATGCGGGAATTTTGTAGTTCTGAAACTATTGTGGAGTTAGTCGAGTATTATT 1174

QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
 DB 1175 GTAGTAGGAAGTATGTAGGA-----GCTTTGTGTCGCCCT----- 1210
 QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120
 DB 1211 GTCCTAGCTGCAGGTATATATCTTTTGGACATTTGTCGGATCTTTTGG---CAGGA 1267
 QY 121 GluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
 DB 1268 TCTGACCTCGCAAACTGTTGGCAGGATTTGTTAAACATCGA-----GGAAG 1315
 QY 141 ProLeuThrGlu-----SerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPhe 157
 DB 1316 CCTATACAGAAATAGATAAAACATAAATTAATGTACTACTTCTATCGTAACACCTATA 1375
 QY 158 ArgGlnIleLeuGlnSerTrpAsnThrAlaLeuAspTrpArgLysLeuLysArgLeu 177
 DB 1376 AAAAATCAACTGTGATAAATATCAAGAAATTTTCGATAAATGGGAGCCAGCACGT----- 1429
 QY 178 GlnAlaProGlyLeuProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIle 197
 DB 1430 -----ACACACGCTAATGCTAAACAGTACATGATCTCTTTACT 1468
 QY 198 ArgPheGluAsnValHisAsn-----AspPheIleArgGluIleProGlyPheGln 214
 DB 1469 ACCTTAGAACCTATAATAGATAAAGATTTAGATATGTTAAATAAATAATGCTAGTATCGA 1528
 QY 215 LeuGluThrTyrlsThrLeuLeuLeuProIleTyrlsAlaGlnAlaAlaAsnPheHisLeu 234
 DB 1529 ATACCACAA-----CTCCCTGCAATATGCACAAATAGTACTTGGCACTTG 1573
 QY 235 AsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSer 254
 DB 1574 AATTTATTAACATCTGCTACTACTATTACATATATG-----CTGCAAAATCAA 1624
 QY 255 GlnIleGluProAsnAlaGlyThrSerAspAspTyrlsLys---LeuLeuLysGluAsn 273
 DB 1625 GGTATAAATCCAAGTACTTCAATTCATCTAATCTATCAGGCTATTTAAACGTTAA 1684
 QY 274 IleProLysTyrlsSerAsnTrpCysAlaAsnThrTyrlsArgThrGlyLeuLysAsnLeuArg 293
 DB 1685 ATACAGAAATATCTGACTATTGTATACAAACCTCAATCAGGACTTAACCTATGATTAGA 1744
 QY 294 AspGluProAsnMetLysTrpSerIlePheAsnAspTrpArgArgTyrlsMetThrIleThr 313
 DB 1745 ACTAATACTAACCCACATGGAATATGATATACTTACCTGCTTAGAATGACTTAACT 1804
 QY 314 ValLeuAspThrIleSerGlnPheSerLeuTyrlsAspIleLysArgTyrlsArgAspSerIle 333
 DB 1805 GTGTTAGATCTTATTGCTATTTTCCAAATTTATGATGACCCAGAAAAATAT----- 1852
 QY 334 GlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrlsThrGlu 353
 DB 1853 -----CCAATAGAGGTAAATCTGAATCTTATCAGAGAAGTTTAT---ACGAAT 1897
 QY 354 IleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrls 373
 DB 1898 GTTAAATTCAGATACATTT-----AGAACCTAAACAGAACTAGAAAAAT 1939
 QY 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrlsThr 393
 DB 1940 GGATTAACATAGA---AATCCTACATATTATTCTTGGATAAACCAAGGGCGTTTTTACACA 1996
 QY 394 GluAsnThrAsn-----PheGlyAsnArgLeu 402
 DB 1997 AGAAATTCGAGACATCTCTTGATCTTATGATATTTTCTTTTACAGGTAAACCATG 2056
 QY 403 ValGlyIleSerAsnArgAspAlaProThrTyrlsSerAsnThrIleThrGluThrLeuTyrls 422
 DB 2057 GCCTTTACATACTAATATGATGAT-----CGCAACATAAATCTGGCAGCGGTTCAT 2107
 QY 423 GlyGluArgThrGlySerProThrThrLysThrIle-----ArgProPhe 437

DB 2108 GGAATATATATTCTCAAGACACATCCAAAGTATTTCTCTTTTATAGAAACAACTATT 2167
 QY 438 GluSerTyrlsValSerIleValThrAspArgGlnSerProProValSerProIleGln 457
 DB 2168 GAT-----AAGTTCGAAATGTGCAGATAGAGTACTCAGATATATAA----- 2212
 QY 458 ProHisPheIleIleAsnGlnIleGluLeuTyrlsLeuAsnGlySerSerAsnAsnThrLeu 477
 DB 2213 -----TATGAAATGATATATTTTTCGAATAGCAGTGAAGTATTT 2251
 QY 478 LysTyrlsSerAlaGlyGlySerLeu---SerAsnTyrlsGlnAsnThrThrPhePheGlnPhe 496
 DB 2252 CGATATTTCATCCCAATTCACAAATAGAAATTAATATAAAGAACTGATTTATATGAT 2311
 QY 497 ProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsn 516
 DB 2312 CCAAAACAAACATCGAAAAAT-----CAAGAA 2338
 QY 517 TyrlsSerHisIleLeuSerHisPheSerLeuPheThrTyrlsTyrlsValIleGlyLeuGln 536
 DB 2339 TATGGTCACTACTCTATCGTATATAAAACATGATAATTATATATTTTCAGTAGTAGAGAA 2398
 QY 537 LeuGlnIleLeuAspThrGlyValLeuGlyTrpThrHisSerSerValAspArgTyrlsAsn 556
 DB 2399 AGAAGAGAGT-----GCATTTAGTTGGACACATACTAGTGTGTGTTCCAAAAT 2449
 QY 557 AlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThr 576
 DB 2450 ACATAGATTAGTATACATCACCCCAATTCACGCTCTAAACACTTTAAAGTAGTAGTCT 2509
 QY 577 AsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrlsLeuGlnSer 596
 DB 2510 AATTCAAAATTTGAAAAGCGCTCGGTACACAAAGTGAATACTTGGTAATTTCTTAAGAT 2569
 QY 597 GlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrlsPheIleArg 616
 DB 2570 AGTATGAATTTTAGAGTTAGATTT---TTAAAAAATGTTTCTCAACAATATCAAGTAGCT 2626
 QY 617 LeuArgTyrlsAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
 DB 2627 ATTCGTTATGCTACTAATGCT-----CCAAAGACACACAGTAGTATCTTAACC 2671
 QY 637 GlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrlsAsn 656
 DB 2672 GGAATAGACTACTATAAGTGTGGAG---CTCCCTAGTACCCTTCCCGCCAAACCCAAAT 2728
 QY 657 -----AsnLeuGlnTyrlsGlyAspPheGlyTyrlsPheGlnPheProSerThrValThrLeu 674
 DB 2729 GCTACAGATTAAACATATGCAGATTTTGGATATGATCAATTTTCAAGAACACAGTTTCA--- 2785
 QY 675 ProLeuAsnArgAsnIle-----ProPheIle-----PheAsnArg 686
 DB 2786 -----AATAAAACATTTGAAGGAGAGACACTTTTATTAATGACCTTTATATGCTACA 2839
 QY 687 AlaAspValSerAsnSerIleLeuIleAspLysIleGluPheIleProIleThrSer 706
 DB 2840 CCAATCATTCATATAATA---TATATTGACAAAATTTGAATTTATTTTCAATCACTCAA 2896
 QY 707 SerMetHisGlnAsnArgGluLysGlnLeuGluThrIleGlnThrLysIleAsnThr 726
 DB 2897 TCTGTATTAGATTATACAGAGAACAAAATATAGAAAAAACACAGAAAAATAGTGAATGAT 2956
 QY 727 PhePheThrAsnHisThrLysThrLeu 735
 DB 2957 TTATTTGTTTAAATAAAACAAAGTTCTT 2983

RESULT 13

AF285775 4959 bp DNA linear BCT 16-AUG-2000
 LOCUS Bacillus thuringiensis subsp. finitimus parasporal inclusion
 DEFINITION protein Cry (cry) gene, complete cds.
 AF285775
 ACCESSION


```

QY 479 TyrSerAlaGlySerLeuSerAsnTyr-----GlnAsnThrThrPhePheGln 495
Db 2349 TATCAAAAAATTTTAAATGTAATCAATCAAAATGAACCTCAAGAAACTACA-----2399
QY 496 PheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsn 515
Db 2400 -----AACTATCTCTAATGATTATGTTGGTGTCA---AAACGCCAA 2435
QY 516 AsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeu 535
Db 2436 AAATTCACACATAATTTATCTCATTTTCCATTAATCATCCAC-----2477
QY 536 GlnLeuGlnIleLeuAsp-----ThrGlyValLeuGlyTrpThrHisSerSer 551
Db 2478 CAAAGTAGAGTTTGCTGAGTATTTTCACTCTATATTTGCAATAGTTGGACACACAATAGT 2537
QY 552 ValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGly 571
Db 2538 GTAAACTCCCAAAATTTAATATCAGAAAGTGTGAGTACACAAATCCCATTTGGTAAAGCT 2597
QY 572 AsnAsnLeuAspThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeu 591
Db 2598 TAGCAAGTTACTACCAATTCA---GTTATAGAGGACCGAGTTTACAGTGAGATTTA 2654
QY 592 ValTyrLeuGlnSerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGln 611
Db 2655 ATAGAACTTCGATGAT-----AAATGTTCTATTAAATGTAAAGCT---AGTTCTTTAAAA 2705
QY 612 SerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIle 631
Db 2706 AAATACGCTATAAGTCTATTTTATGCTGCATAAATACGCAATAGCTGATCAATAGACGTA 2765
QY 632 SerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnThrPheSer 651
Db 2766 GTGATTCGAGACGAGGATTCTTA-----TTGCAACTACTCTTTTCT 2807
QY 652 GlyThrAsnTyrAsnAsn-----LeuGlnTyrGlyAspPheGlyTyrPhe 666
Db 2808 AGAAAAAGGAACAATAATTTTACAATTCAGACCTTAACTATAAGGATTTCAATATCAT 2867
QY 667 GlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArg 686
Db 2868 ACACATTTTAGTTGATATTGAATTTACCGAAAGTGAAGAAATTCATATCCATTTTGAAGCGA 2927
QY 687 AlaAsp-----ValSerAsnSerIleLeuIleAspLysIleGluPheIleProIle 704
Db 2928 GAGGATGATTATGAGGAGGAGTGATTTTAAATGATAAATAGAGTTCAACCTATA 2987
QY 705 ThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLysIle 724
Db 2988 GATGAAATATAT-----ACTAATGAATGAATTTAGAGNAGGCAAGAACGAGTG 3038
QY 725 AenThrPhePheThrAsnHisThrLysThrLeu 735
Db 3039 AATGTATTATTATAACGCAACAAACGCTTTG 3071

RESULT 14
LOCUS BTH251977
DEFINITION Bacillus thuringiensis subsp. medellin cry29Aa gene for Cry29Aa protein.
ACCESSION AJ251977
VERSION AJ251977.1 GI:17977978
KEYWORDS Cry29Aa gene; Cry29Aa protein.
SOURCE Bacillus thuringiensis serovar medellin
ORGANISM Bacillus thuringiensis serovar medellin
          Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE 1
AUTHORS Delecluse,A. and Orduz,S.
TITLE Characterization of two new mosquitocidal toxins, Cry29A and Cry30A, from Bacillus thuringiensis medellin

```

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
2 (bases 1 to 6009)
Delecluse,A.
Direct Submission
Submitted (22-DEC-1999) Delecluse A., Bacteries & Champignons
Entomopathogenes, Institut Pasteur, 25, rue du Dr Roux, 75724 Paris
Cedex 15, FRANCE

FEATURES
source

1. 6009

/organism="Bacillus thuringiensis serovar medellin"
/mol_type="genomic DNA"

/isolates="163-131"

/sub_species="medellin"

/db_xref="taxon:79672"

/country="Colombia"

642. 2594

/gene="cry29Aa"

642. 2594

/genes="cry29Aa"

/function="putative mosquitocidal toxin"

/codon_start=1

/evidence=experimental

/transl_table=11

/product="Cry29Aa protein"

/protein_id="CAC80985.1"

/db_xref="GI:17977979"

/db_xref="GOA:Q8VNX2"

/db_xref="UniProt/TrEMBL:O8VNX2"

/translation="MNSFONKVEYLINAPSSINIPNNYSKYPIANTILNQTOMNTNY
KWTNWCADNNLKSINPPFANLQNSLVGLFALTAALASLSAPITGTSIAAGTAIA
AAIPLPSPQNNLPDLALISEATLYSFQDQREDALTRLESKDSVKVFNFAFT
FWINPNSTNTTTFRRFQVNGFVGSMAFFRAKNEYPIILSTYAQARLHLRLD
GITYASKWLSROGDMPGLDYLFKNKYCNEYIEHCIKWNSLSLLKSGANWLEY
NOVYRTLASVLVDVLSLFSYDRLYKERLSVEILTKLYTDPINYHRTGSLADESK
YTLPEPLTQLYTLTPYSNIFYNMGHTWYRILSPDKIFAERSFGQSSVIDKVPVI
PNDKSLIYKIRAYDNHGLFNVMYFGWGEKQIQIKIGSSSEIYIKNCKTRLADV
ISHDLDEKKCYSAFWSTTISLENEIKNDIITQIPAVKAYQLGVSQVTKPGPTGG
NINLKSNDYLRISCOHLSNVTKYFVIRYATNGSLNTRPIINITPGMTPQGMVLD
NTPSGTGYNLKYNFGYKFLKEFLVLPNQISLTLNRSQDQNSNLSILLDRIFLPI
TPSIRKSEQNLEKNQKTVNKLFFN"

ORIGIN

Alignment Scores:

Prod. No.: 5.01e-54 Length: 6009
Score: 850.00 Matches: 232
Percent Similarity: 46.97% Conservative: 124
Best Local Similarity: 30.61% Mismatches: 276
Query Match: 21.97% Indels: 126
DB: 1 Gaps: 22

US-10-783-417-2 (1-735) x BTH251977 (1-6009)

QY 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 642 ATGAATTCATTTTCAAAATAAAATGAATATGAATAATATAATGCTCCATCGAGT-----695
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 696 ATTTCTTAATATACCTAATAATATTCTTAAATATCCAAATAGCAATGCAATGATAATAAATCAAACT 755
QY 41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCys-----GlnGlyAsnThrGln 58
Db 756 ATGCAAAACACCAATTTACAAAGATGAGACTAATATGTCGACTAATGATAATAAATCAAA 815
QY 59 TyrGlyAspAsnPheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGly 78
Db 816 TCTATAAATCCTTTTGAAGCTAATCTTCAAAATAGTTTAGTAGTCTTTTGTCTATAACT 875
QY 79 ThrIleValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyPro 98
Db 876 GCTGCTATTGCTTCATTGTTGTGAGTCCCAATTAAGTGGTGTACATCTATAGACAGCTGGC 935
QY 99 IleGlyIleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpPro 118


```

Db      936 ACAGCTATAGCAGCAGCTATA-----ATACCTATACTTTGGCCT 974
Qy      119 AlaGlyGluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheVal 138
Db      975 TCT-----CAAGAAATAATACTCTCCGATAAACTTCTAGCTATTTCAGAACCAACCCTT 1028
Qy      139 AspThrProLeuThrGluSerIleLysGlnLeuLysLeuThrGlnLeuGluGlyPheArg 158
Db      1029 TATTCATTTCAAGATCAAGGGTTAGAGAGATGCACCTTACGCCGATTAGAAAGTTTAAAA 1088
Qy      159 GlnIleLeuGlnSerTyAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGln 178
Db      1089 GATAGTGTAAATATTTTGGAAATGCCTTCACATTTTGGATAAT----- 1133
Qy      179 AlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArg 198
Db      1134 -----AATCCTAATTTCAACC-----AACCAACTACTGTCAGAGAAAGA 1172
Qy      199 PheGluAsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTy 218
Db      1173 TTCAGAGGTTTAATGTAGATTGTAGGCAGCATGGCCTTCTTTAGAGCAAAAGCACTAT 1232
Qy      219 LysThrLeuLeuLeuProIleTyAlaGlnAlaAlaAsnPheHisLeuAsnLeuGln 238
Db      1233 GAACCAATACTGTATCACTATGACACAGCTGCAAGGTGTACATTTACTTCTATTACGG 1292
Qy      239 GlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluPro 258
Db      1293 GATGGAATTACGTATGCTGAAAAATGGAATTTAAGTAGACAAAGGGATGACATG----- 1346
Qy      259 AsnAlaGlyThrSerAspAspTyTyLysLeuLeuLysGluAsnIleProLysTySer 278
Db      1347 -----CCCGGGGATTTACTTTTACAAAGAATTTAACAAGATTTGCAATGATATATT 1397
Qy      279 AsnTyTyCysAlaAsnThrTyArgThrGlyLeuLysAsnLeuArgAspGluProAsnMet 298
Db      1398 GACATTTGTATAAGTGTACATGAAGTTTGTCTCTCTTAA-----AGTGTAGGTGCA 1454
Qy      299 LysTrpSerIlePheAsnAspTyArgTyMetThrIleThrValLeuAspThrIle 318
Db      1455 AATTTGGCTAGAAATACAAATCAATCGTACATTTCTTAACAGCCTCTGTATTAGATGTCATT 1514
Qy      319 SerGlnPheSerLeuTyAspIleLysArgTyArgAspSerIleGlyIleGluVal 338
Db      1515 TCTCTTTTTTCATCATATGACCCCGTTTATATAAAGAAAGATTA---AGTGTGAGATT 1571
Qy      339 LysGlyIleLysAsnGluLeuThrArgGluIleTyThrThrGluIleAsnPheAspArg 358
Db      1572 -----CTCAAGAAATAATTATATACAGATCCAAATTAATTATCATAGA 1613
Qy      359 LeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyAsnLeuThrArgAla 378
Db      1614 GGCATTTTCAAGACAGATGAAGCAATATACACTTGAGCAACTTTGTTTACTCAA 1673
Qy      379 SerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyThrGluAsnThrAsnPhe 398
Db      1674 CTATATACACTAATCTTTATAGTAATATATTTTATAAATTACATGGGACACACCAAT--- 1730
Qy      399 GlyAsnArgLeuValGlyIleSerAsnArgAspAlaProThrTySerAsnThrIleThr 418
Db      1731 -----ACTTATAGATATTATCACCG 1751
Qy      419 GluThrLeuTyGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGlu 438
Db      1752 GACAAAATTTTTCAGAGCGTCTCTTTGGAAACAAGACGCTATATAGAT----- 1802
Qy      439 SerTyLysValSerIleValThrAspArgGlnSerProValSerProIleGlnPro 458
Db      1803 -----AAAGTCCCTGTAATCTCTACGATAATCA----- 1832
Qy      459 HisPheIleLeuAsnGlnIleGluLeuTyLeuAsn----- 470
```

```

Db      1833 -----ATTATTATAAAATTCGACCATATGATAATCATTAATGGTTTATTATTATGTAATG 1886
Qy      471 -----GlySerSerAsnAsnThrLeuLysTySerAlaGlyGlySer 484
Db      1887 TACTTTGGTGTTTTGGGACGGAGAAAGATCAATACAAATAATAATTCGAGGTAGTTCT 1946
Qy      485 LeuSerAsnTyGlnAsnThrThrPhePheGlnPheProArgLysLysAspCysAsnLeu 504
Db      1947 ACAGAAATATAC-----ATTAATAATTTGTACT--- 1973
Qy      505 ValIleAspProGlyCysSerProAsnPheAsnAsnTySerHisIleLeuSerHis--- 523
Db      1974 -----CATCGCTTAGCAGATGTCATCTCACATGAT 2003
Qy      524 -----PheSerLeuPheThrTySerTyValIleGlyLeuGlnLeuIleLeuAsp 541
Db      2004 CTTCATGAAAAAACAATGTTATTTCTTT----- 2033
Qy      542 ThrGlyValLeuGlyTrpThrHisSerSerValAspArgTyAsnAlaIleSerAspLys 561
Db      2034 -----GCATGGACATCTACTCAATTTCTTCTTGAATAATGAATAAATAATGAT 2081
Qy      562 IleIleThrMetIleProAlaIleLysGlyAsnLeuAspThrAsnSerLysValIle 581
Db      2082 ATTATCACACAATCCCTGCTGTAAAGCCTTCAATTAGGTGTGCAATCACAGTGATT 2141
Qy      582 GluGlyProGlyHisThrGlyGlyAsnLeuValTyLeuGlnSerGlnGlyArgLeuGlu 601
Db      2142 AAGGTCAGGTTCATCTACTGTTGGAGACTTAATAATTTAAATCAATGATTATTACGT 2201
Qy      602 IleThrCysGlu---ThrProAsnSerThrGlnSerTyPheIleArgLeuArgTyAla 620
Db      2202 ATATCATGTCAACACTTATCTAATGTTACAAAAAATATTTGTAAAGATTCGTACGCT 2261
Qy      621 ThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGly 640
Db      2262 ACAATGGTTCTCTAAATACTAGGCCCATTTAAATATTACAAATTTCA-----GGA 2312
Qy      641 IleProGlnArg-----LeuAsnAsnThrPheSerGlyThrAsnTyAsnAsnLeu 658
Db      2313 ATGACACCACCAAGGTATGGTACTCGATTAATCTTTCTGGTACAGTTATAGCAATTTA 2372
Qy      659 GlnTyGlyAspPheGlyTyPheGlnPheProSerThrValThrLeuProLeuAsnArg 678
Db      2373 GAATATCAAAATTTTGGCTATAAGAAATTTCTAAAGAGGTAAACATTAATCCAAATCAA 2432
Qy      679 AsnIleProPheIlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleAsp 697
Db      2433 AGTATATCTCTTACATTAATTAATCGTTCTGATCAAAATTCAAATTTTACTCTCTGAT 2492
Qy      698 LysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeu 717
Db      2493 AGAATTGATTTCTACCAATTTACTCCCTCTATACGAAAAAGTAAAGAACACAAATTTA 2552
Qy      718 GluThrIleGlnThrLysIleAsnThrPhePheThrAsnHisThrLysThrLeu 735
Db      2553 GAAAAAATCAGAAAAACGGTAAACAAGTTATTTTCAATTAATAATAAAGGGTTG 2606

RESULT 15
AF132928 4896 bp DNA linear BCT 14-JAN-2000
LOCUS Bacillus thuringiensis subsp. finitimus Cry28Aa1 delta-endotoxin
DEFINITION gene, complete cds.
ACCESSION AF132928
VERSION AF132928.1 GI:4574729
KEYWORDS Bacillus thuringiensis serovar finitimus
SOURCE Bacillus thuringiensis serovar finitimus
ORGANISM Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE 1 (bases 1 to 4896)
AUTHORS Wojcicheowska,J.A., Lewitin,E., Revina,L.P., Zalunin,I.A. and Chestukhina,G.G.
```


TITLE Two novel delta-endotoxin gene families cry26 and cry28 from
 JOURNAL *Bacillus thuringiensis* ssp. *finitimus*
 PUBMED FEBS Lett. 453 (1-2), 46-48 (1999)
 REFERENCE 10403372
 2 (bases 1 to 4896)
 AUTHORS Wojciechowska, J.A., Lewitin, E.I. and Chestukhina, G.G.
 TITLE Direct Submission
 JOURNAL Submitted (04-MAR-1999) Laboratory of Protein Chemistry, Institute
 of Microbial Genetics, Dorozhny proezd, Moscow 113545, Russia
 Location/Qualifiers

FEATURES
 source

1. 4896
 /organism="Bacillus thuringiensis serovar *finitimus*"
 /mol_type="genomic DNA"
 /strain="VKPM B-1161"
 /sub_species="finitimus"
 /db_xref="taxon:29337"
 1125..4458

CDS

/codon_start=1
 /transl_table=11
 /product="Cry28Aa1 delta-endotoxin"
 /protein_id="AAD24189.1"
 /db_xref="GI:4574730"
 /translation="MAQYKIKGVQSTVNSSEIFNPEVDSSTVAVUSAGIVVGT
 ILTAPASVNPQGVLIISGLTAPVLPDPPEPKKIQSQFMKGHSDLLNQTISTAVKE
 IALAHNGFKDVLITYERAFNDKRNPSANTARLVSRPENAHFNVFNSPOLQPTY
 DTLSSCTEAAHLNLHOGVQADQWADPHSPMLKSSGTYYDELLVYIKYIN
 YCTKYHKLHKESEKTDAYNTYREMTLIVLDLVATPPYDIRRPPRGVLEL
 TREVTSLDLTPPLGLTWLSDIELYTESVAEGDYLGLRESKYVTGNQPTMKNIY
 GNTNRLSQLITLPGEFWTHLSINRPFQTIAGINKLYSLIOKIIVPTFPKNDNEVQKN
 FNVNPNQPEITLTPNGGNSQKRNLSHPFLIIHKLFASYFHSIFALGWTHN
 SVNSQLISVESVTOIPLVAYEVNNSVIRGPTGGDLIELDKSCIKKASLKK
 YALSIFYANNAIVADVDSGAGVLOPTFSRGNNTFTQDLNLYKDFVHTLLVD
 IELPESHIHLKREDDVEGVILLIDKLEFKIDENYENMLEKAKAVNVLFI
 ATNALKMDVTDHIDQANLVFCISDILYAKEKIKETPCIKFAKOLSOARNLSDPNF
 NNLNENSWANTGVTIIEGDLYGRALQLSAARDENPPYLYOKIDESLKPTRY
 QLRGVEGSDLELVRGADTIIVNVPGDLEILSYSAPIPCBEIETRLTTCGAL
 DRCKQSVNSAADYRPPQVNGDPHAFSGHIDTIDNNNLGIWIIKFIATPDGAT
 FGNLELIELGSLGALQVROKQKGNKTQKEEAAKLYAAAKQINQLFADSQG
 TKLPDFTFNSLADLVYKIRDVYSEVLVPIGLNYDLFWELNRIQNALDYDAR
 NTVNGEPRNGLANWASNTVEROIQAHPCWYSLGNAQVAQSINVPDHCYVLRVT
 AKGIGNGYVTILCANHIDLTFTSSCDSGFTTSSNELAAVYTKLEIFPPTDQIRI
 EIGETRSTFYVESVDLIRMED"

ORIGIN

Alignment Scores:
 Pred. No.: 4,62e-54 Length: 4896
 Score: 849.00 Matches: 229
 Percent Similarity: 48.76% Conservative: 106
 Best Local Similarity: 33.33% Mismatches: 248
 Query Match: 21.94% Indels: 104
 DB: 1 Gaps: 25

US-10-783-417-2 (1-735) x AF132928 (1-4896)

QY 68 SerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAla 87
 DB 1210 AGCAGTGATACAGTCGCTAGTAGAAGCGAGGAGTTGTTAGTGTGGTACTACTACTGACA 1269
 QY 88 GlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIleIle 107
 DB 1270 -----GCCTTTGCATCATTTGTTAATCCA-----GGTGTGGTACTTATA 1308
 QY 108 SerPheGlyThrLeuIleThrValPheThrProAlaGlyGluGlnAsp---LysThrVal 126
 DB 1309 TCATTGGAACTTGGCTCCCGTCTCTTTGGCTGTATCCAGAGGAAGATCAAAAAAAT 1368
 QY 127 TrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeuThrGluSerIle 146
 DB 1369 TGGTCACATTTATGAAACACGAGAGAGACCTTTTAAATCAACAATTTCTACAGCTGTA 1428
 QY 147 LysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGlnSerThrAsnThr 166
 DB 1429 AAGAAATAGCATAGCTCATCTAAATGTTTAAAGATGTTAATACGTACTATGAAGA 1488

QY 167 AlaLeuAspAspThrArgLysLeuLysArgLeuGlnAlaProGlyLeuProSerSer 186
 DB 1489 GCATTTAATGATTGGAAGA-----AATCCAAGTGCA 1521
 QY 187 AlaLeuGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnApphe 206
 DB 1522 -----AATAGTCCAGATGGGTATCACAGAGATTTGAAACGCTCATTTCAATTTT 1572
 QY 207 IleArgGluIleProGlyPheGlnLeuGluThrThrLysThrLeuLeuProIleTyr 226
 DB 1573 GTAAGCAATATGCCACAACCTCCAACTTCCACGTATGACACATATTATTAAAGTGTCTAT 1632
 QY 227 AlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnAlaGluLeuAlaAspGlu 246
 DB 1633 ACAGAAAGCTGCAATTTTACATTTGAATTTATTACATCAAGGTGTACAATTCGCGGATCAA 1692
 QY 247 TrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTyr 266
 DB 1693 TGGAAATGCAGATCAACACCATTTCCCAATTTGTAAGTCTCATAGGTACT-----TAT 1743
 QY 267 TyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyrArg 286
 DB 1744 TATGACGACCTATTGGTATATTATTTGAAAGTATATTAAATTATGACCAAGACATACCAT 1803
 QY 287 ThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTyr 306
 DB 1804 AAAGGATTGAATCACCTTAAAGAAATCAGAAAAATCACAATGGGATGCTTATATAACATAT 1863
 QY 307 ArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIle 326
 DB 1864 CGTCGAGAAATGACCTTAATTTGTTGGATCTTGTGCGCACTTTCTTTTATGATATA 1923
 QY 327 LysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThr 346
 DB 1924 CCGTGGTTT-----CCAGGAGGAGTAGAAGTAACTAGAAATTAACA 1959
 QY 347 ArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeuArgValGlnPro 366
 DB 1960 ACAGAGGTTTATACAAGTTTAGATCATTTTAAACACGACCA----- 2001
 QY 367 AsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeu 386
 DB 2002 -----GGGCTATTACTTGGCTG 2019
 QY 387 GluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIleSer 406
 DB 2020 TCAGATATTGAGTTATACAGGAGAGTGTCGAGAGGCGATTTATTCAGGTATTT--- 2076
 QY 407 AsnArgAspAlaProThrTyrSerAsnThr-----IleThrGluThrLeuTyrGly 423
 DB 2077 ---CGAGAGCTAAATATTATCTGCTAATCAATTTTACGATGAAAAATATTTATGCT 2133
 QY 424 GluArgThrGly---SerProThrThrLysThrIleArgProPheGluSerTyrLysVal 442
 DB 2134 AATACAAATAGATTAAAGTAAGCAGCTCATTTACATTTATCCAGCGCAA---TTTATGACT 2190
 QY 443 SerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIleIle 462
 DB 2191 CACTTAAGCATAAACCGCTCTTTTCAACAATAGCTGGTATAAATAGTTATACAGTTTA 2250
 QY 463 AsnGlnIleGluLeuThrLeuAsnGlySerSerAsnAsnThrLeuLysTyrSerAlaGly 482
 DB 2251 ATTCAAAAAATCGTATTACAACTTTTAAAAACGATAAT-----GAATATCAAAAAAT 2304
 QY 483 GlySerLeuSerAsnTyr-----GlnAsnThrThrPhePheGlnPheProArgLys 499
 DB 2305 TTTAATGTGAATTAATCAAAATGAACCTCAAGAACTACA----- 2343
 QY 500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHis 519
 DB 2344 -----AATATCTCTAATGATTATGGTGGTTCA---AACGCCAAAAAATTCACACAT 2391

Qy	520	IleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIle	539
Db	2392	AATTATCTCATTTTCCATTAATCATCCAC	2433
Qy	540	LeuAsp-----ThrGlyValLeuGlyTrpThrHisSerSerValAspAsgTyr	555
Db	2434	GCTGAGTATTTTCACTCTATATTGCAATAGTTGGACACACAAATAGTGAACCTCCCAA	2493
Qy	556	AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp	575
Db	2494	AATTTAATATACAGAAAGTGTGACGACACAAATCCATGGTTAAAGCTTACGAAATT---	2550
Qy	576	ThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln	595
Db	2551	ACTAACCAATTTCAGTTATAAGAGGACCAAGGTTTACAGGTGGAGATTTAATAGAACTTCGT	2610
Qy	596	SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIle	615
Db	2611	GAT-----AAATGTTCTATTAAATGTAAAGCT---AGTTCTTTAAAAAATAACGCTATATA	2661
Qy	616	ArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIle	635
Db	2662	AGTCTATTTTATGCTGCAATAACGCAATAGCTGTATCAATAGACGTAGTGATTCGCGA	2721
Qy	636	ProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyr	655
Db	2722	GCAGGAGTTCTA-----TTGCAACCTACCTTTCTAGAAAAGGGAAC	2763
Qy	656	AsnAsn-----LeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSer	670
Db	2764	AATAAATTTTACAATTTCAAGACCTTAACATAAAGGATTTTCAATATCATACACTTTTAGTT	2823
Qy	671	ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAsp-----	688
Db	2824	GATATGAAATTTACCCGAAGTAGAAGAAATTCATATCCATTTGAAGCGAGAGGTGATAT	2883
Qy	689	ValSerAsnSerIleLeuIleAspLysIleGluPheIleProIleThrSerSerMet	708
Db	2884	GAGGAGGAGTGATTCITTTTAATGTAAATATAGAGTTCAAACTATAGATGAATAATAT	2943
Qy	709	HisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePhe	728
Db	2944	-----ACTAATGAATGAATTTAGACGAGCAAGAAAGACAGTGAATGTATTATT	2994
Qy	729	ThrAsnHisThrLysThrLeu	735
Db	2995	ATAAACGCAACAAACGCTTTG	3015

Search completed: December 12, 2005, 02:22:23
Job time : 9851 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 11, 2005, 20:13:22 ; Search time 1089 Seconds

(without alignments)
4498.209 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869

Sequence: 1 MNQNDNNEVEIIDSHTSPY.....KLETIQTKINTFFNHTKTL 735

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10783417/runat_01122005_141956_24301/app.query.fasta_1.903
-DB=N Geneseq -QFMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40 cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10783417 @CGN 1 1 752 @runat_01122005_141956_24301 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 21.*

1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3869	100.0	2208	13	ADR89394
2	3314	85.7	2235	13	ADR89396
3	3104.5	80.2	2085	13	ADR89398
4	1061.5	27.4	4571	1	AAN93059
					ADR89399 Delta-end

5	1060.5	27.4	3543	2	AAQ14669
6	1060.5	27.4	3543	2	AAQ81178
7	1059.5	27.4	3940	1	AAN93054
8	1059.5	27.4	4934	1	AAN81490
9	1032.5	26.7	2130	12	ADP71294
10	1032.5	26.7	2246	12	ADP71295
11	899.5	23.2	2061	2	AAQ14670
12	899.5	23.2	2061	2	AAQ81180
13	881	22.8	3756	1	AAN50525
14	849	21.9	4896	6	AAD43974
15	849	21.9	4896	10	ADF31301
16	849	21.9	4896	10	ADF31306
17	775.5	20.0	3507	2	AAT43321
18	759	19.6	3684	6	ABK14949
19	759	19.6	3684	14	AEA61392
20	758	19.6	3684	1	AAN82106
21	757.5	19.6	3535	1	AAN91003
22	747.5	19.3	4184	1	AAN96139
23	711.5	18.4	4065	1	AAN92515
24	705	18.2	4391	14	AEA61394
25	703	18.2	3633	6	ABK87235
26	703	18.2	3633	12	ADL15306
27	703	18.2	3633	14	AEA45608
28	703	18.2	6613	6	ABK87248
29	703	18.2	6613	12	ADL15321
30	703	18.2	6613	14	AEA45623
31	696.5	18.0	10974	1	AAN60055
32	689.5	17.8	4188	8	ABX13487
33	689.5	17.8	4188	14	ABE90816
34	678.5	17.5	2003	6	ABK87237
35	678.5	17.5	2022	6	ABK87242
36	678.5	17.5	2022	12	ADL15316
37	678.5	17.5	2022	14	AEA45618
38	677	17.5	2478	12	ADP43079
39	676	17.5	1953	3	AAQ43223
40	676	17.5	1953	3	AAQ09499
41	676	17.5	2425	2	AAQ05679
42	675	17.4	3621	6	ABK87234
43	675	17.4	3621	12	ADL15304
44	675	17.4	3621	14	AEA45606
45	675	17.4	3621	14	AEA34684

ALIGNMENTS

RESULT 1	
ADR89394	
ID	ADR89394 standard; cDNA; 2208 BP.
XX	
AC	ADR89394;
XX	
DT	18-NOV-2004 (first entry)
XX	
DE	AXMI-006 coding sequence.
XX	
KW	ss; gene; delta-endotoxin; delta-endotoxin associate polypeptide;
KW	expression cassette; transformation; transgenic; plant; bacteria;
KW	lepidoptera; coleoptera; pest; pesticide; resistance;
KW	pesticidal activity.
OS	Bacillus thuringiensis.
XX	
FH	Key Location/Qualifiers
FT	CDS
FT	1..2208
FT	/*tag= a
FT	/product= "AXMI-006"
XX	
PN	WO2004074462-A2.
XX	
PD	02-SEP-2004.
XX	
PP	20-FEB-2004; 2004WO-US005829.
XX	

PR 20-FEB-2003; 2003US-0448632P.
PR 20-FEB-2003; 2003US-0448633P.
PR 20-FEB-2003; 2003US-0448797P.
PR 20-FEB-2003; 2003US-0448806P.
PR 20-FEB-2003; 2003US-0448810P.
PR 20-FEB-2003; 2003US-0448812P.
PR 19-FEB-2004; 2004US-00781979.
PR 19-FEB-2004; 2004US-00782020.
PR 19-FEB-2004; 2004US-00782096.
PR 19-FEB-2004; 2004US-00782141.
PR 19-FEB-2004; 2004US-00782570.
PR 19-FEB-2004; 2004US-00783417.
XX (ATHE-) ATHENIX CORP.
XX Carozzi N, Hargiss T, Koziel MG, Duck NB, Carr B;
PI WPI; 2004-635574/61.
DR P-PSDB; ADR89395.
XX New isolated delta-endotoxin and delta-endotoxin-associated nucleic acids
PT and polypeptides, useful for killing lepidopteran or coleopteran pests or
PT for producing organisms with pesticide resistance.
XX Claim 1; SEQ ID NO 6; 178pp; English.
XX This sequence encodes an isolated delta-endotoxin. Some of the delta-
CC endotoxin coding sequences of the invention have alternative start
CC codons, producing more than one protein from a single open reading frame.
CC The nucleic acid sequences of the invention are useful in DNA constructs
CC or expression cassettes for transformation and expression in plants and
CC bacteria. The nucleic acids and corresponding polypeptides are useful for
CC killing lepidopteran or coleopteran pests. Compositions containing the
CC delta-endotoxins of the invention, and methods for their production, are
CC useful for the production of organisms with pesticide resistance,
CC specifically bacteria and plants. These organisms are useful for
CC generating altered or improved delta-endotoxin or delta-endotoxin-
CC associated proteins that have pesticidal activity, or for detecting the
CC presence of delta-endotoxin or delta-endotoxin-associated proteins or
CC nucleic acids in products or organisms.
XX SQ Sequence 2208 BP; 842 A; 377 C; 308 G; 681 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 2208
Score: 3869.00 Matches: 735
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-783-417-2 (1-735) x ADR89394 (1-2208)

QY 1 MetAsnGlnAsnAspAsnAsnGluTyrGluIleAspSerHisThrSerProTyr 20
DB 1 ATGAAACAAAATAACGATAATAACGAATATGAAATATTGATTCGCATACCTCACCTTAT 60
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
DB 61 TTTCGGAACAGAAACAGTAAATGATTCAGATACCTTACACAAATAATCCAAATCAACCA 120
QY 41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
DB 121 TTACAAACACAAATTACAAAGAGTGCTCAATATGTGTCAAGGAAATACACAAATGGT 180
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80
DB 181 GATAATTCGAGACATTTCGTAGTGTGATACAAATTCCTCAGTTAGTCAGGTACTATT 240
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
DB 241 GTATCCGGTACTCTGTAGCCGGTATAGTGGGCTCATTCTATATCCGACCGATAGGA 300

QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProIleGly 120
DB 301 ATAATAGTGCTATAATAATATCTTTTGGTACCCCTAATCACTGCTTTTGGCCCGGGA 360
QY 121 GluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
DB 361 GAACNAGACAAACAGTATGGACAAATTTATTAATATGGAGAAATTTTGTGTATACA 420
QY 141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
DB 421 CCGTTAACAGAAAGCATAAACAGCTAAAGTTTACAACTTTTGAAGGATTTAGACAAATA 480
QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaPro 180
DB 481 TTCAAAAGCTATATATACAGCATTAGATGATGGAGAAAATTAAGAAAGACTCAAGCTCCT 540
QY 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGlu 200
DB 541 GGATTACCACCATCATCAGCATTAACAACAGCTGCTTGAAGCTTTAAATACGATTTGAG 600
QY 201 AsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThr 220
DB 601 AATGTTCAATGATTTTATTCGAGAAATACCTGCTTCCAACTTGAACCTTATAAAGC 660
QY 221 LeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGly 240
DB 661 CTATTACTACTATTTATGCGCAAGCTGCTAAATTTTCAATTTAAATTTTATTAACAAGGT 720
QY 241 AlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAla 260
DB 721 GCTGAATGCTGATGAATGGAATGCAGATATACCTCTTCAAAATTTGAACCTTAATGCT 780
QY 261 GlyThrSerAspAspTyrTyrLysLeuLysGluAsnIleProLysTyrSerAsnTyr 280
DB 781 GGAACATCAGATGACTATTAATTAATTTTAAAGAAATATACCTTAATATAGTAACCTAT 840
QY 281 CysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTyr 300
DB 841 TGTGCAAAATACCTATAGAACAGGACTAAAAAATCTTAGAGACGAAACAAATATGAAATGG 900
QY 301 SerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGln 320
DB 901 AGTATATTTTAAATGACTATCGAAGATATATGACCAATTAATGATTAGATACCTCTCTCA 960
QY 321 PheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyIleGluValLysGly 340
DB 961 TTTTCTTTATATGATATAAAAGATATAGAGATTCATAGGAGGAATAGAGATAAAGGC 1020
QY 341 IleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuPro 360
DB 1021 ATTAAGAATGAACCTCACAGAGAAATTTATACAACTGAAATATAATTTTGAATCGTCTCT 1080
QY 361 GlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPhe 380
DB 1081 CAACCTTAGAGTTTCAACCCCAATCTAGCTACGATGGAATATAATTTAACAACGTCGCAAGTTT 1140
QY 381 LysLeuPheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsn 400
DB 1141 AAATATTTTCAATTTTATAGAACAAATTTATTTTATACAGAAATATCAAAATTCGGGAA 1200
QY 401 ArgLeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
DB 1201 CGTTTAGTGGTATTTCTAATCGTGATGCACCTACTTATAGCAATACTATAACTGAACT 1260
QY 421 LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440
DB 1261 TTATATCGAGAAAGAACAGGTTCAACCCCAACAAATAAAGACCAATTTGAATCTTAT 1320
QY 441 LysValSerIleValThrAsnArgGlnSerProProValSerProIleGlnProHisPhe 460
DB 1321 AAAGTTTCAATTTGTAATGATAGACAATCACCTCTGTTTCCCTATTTCAACACACTTT 1380
QY 461 IleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsnThrLeuLysTyrSer 480

```
Db 1381 ATAATAATCAAAATTTGAACCTTTATTTAAATGGCTCATCTAACACACATCAATATTC 1440
Qy 481 AlaGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArgGlyLys 500
Db 1441 GCAGGAGGGCTTTATCTAATATTAATCAAAACACAACTTTTTCATTTTCTAGAAAAA 1500
Qy 501 AspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisTle 520
Db 1501 GACTGCATCTAGTTATGATCCAGGTGTTCCAAACCTTTAATACTATAGTCATATT 1560
Qy 521 LeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuIleLeu 540
Db 1561 TTATCCCATTTTTCATTTATTTACTTATTCCTATGTGATGGATTACAGCTACAAATATTA 1620
Qy 541 AspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerAsp 560
Db 1621 GATACAGGTGTATTAGGATGCACACAGTAGTGTGTGATAGATATAATGCAATATCAGAT 1680
Qy 561 LysIleIleThrMetIleProAlaIleIleGlyAsnAsnLeuAspThrAsnSerLysVal 580
Db 1681 AAAATAATTAACAATGATCCAGCAATCAAGGTAACTTGTATACAACTCTAAGGTA 1740
Qy 581 IleGlyGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeu 600
Db 1741 ATTTAGGAGCTGGTCTATACAGGAGAACTTGGTTTATTTTACAAAGTCAAGGCGTTTA 1800
Qy 601 GluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAla 620
Db 1801 GAAATTCATGTGAAACTCTTAATCTTACACATCTTATTTCATTAGCTCGATATGCT 1860
Qy 621 ThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGly 640
Db 1861 ACAATGGTCTCGAAATACTCTTCCTAATATATCTCTTCAATATCCAGGATTAATAGGA 1920
Qy 641 IleProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyr 660
Db 1921 ATACCACCTCAAGACTCAACAAACACTTTTCTGGTACAAATTAATAATTTACAAATAC 1980
Qy 661 GlyAspPheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIle 680
Db 1981 GGAGATTTGGGTATTTCCATTTCCAGTACAGTAACTTACCTTTAATCGAAACATA 2040
Qy 681 ProPheIlePheAsnArgAlaAspValSerAsnSerIleIleIleAspLysIleGlu 700
Db 2041 CCATTTATATTTAATCGTCGAGATGTATCAAAATTCATTTTAAATCAATGATAAAATGAA 2100
Qy 701 PheIleProIleThrSerSerMetHisGlnAsnArgGlyLysGlnLysLeuGluThrIle 720
Db 2101 TTTTATACCAATTTACTTCTCTATGCACCAAAATAGAGAAAAACAAAATTAGAACTATC 2160
Qy 721 GlnThrLysIleAsnThrPhePheThrAsnHisThrLysThrLeu 735
Db 2161 CAAACAAAAATAAATACATTTTTCACAAATACATACAAAAACATTT 2205
```

RESULT 2

ADR89396

ID ADR89396 standard; cDNA; 2235 BP.

XX AC

XX ADR89396;

DT 18-NOV-2004 (first entry)

XX DE

XX AXMI-007 coding sequence.

XX KW

KW ss; gene; delta-endotoxin; delta-endotoxin associate polypeptide;

KW expression cassette; transformation; transgenic; plant; bacteria;

KW lepidoptera; coleoptera; pest; pesticide; resistance;

KW pest

KW pest

KW pest

KW pest

KW pest

KW pest

KW pest

KW pest

KW pest

KW pest

FT CDS

```
FT 1. .2235
FT /*tag= a
FT /product= "AXMI-007"
FT /transl_except= pos:1. .3, aa:Met
```

XX XX

XX WO2004074462-A2.

XX XX

XX 02-SEP-2004.

XX XX

XX 20-FEB-2004; 2004WO-US005829.

XX XX

XX 20-FEB-2003; 2003US-0448632P.

XX PR

XX 20-FEB-2003; 2003US-0448633P.

XX PR

XX 20-FEB-2003; 2003US-0448797P.

XX PR

XX 20-FEB-2003; 2003US-0448806P.

XX PR

XX 20-FEB-2003; 2003US-0448810P.

XX PR

XX 20-FEB-2003; 2003US-0448812P.

XX PR

XX 19-FEB-2004; 2004US-00781979.

XX PR

XX 19-FEB-2004; 2004US-00782020.

XX PR

XX 19-FEB-2004; 2004US-00782096.

XX PR

XX 19-FEB-2004; 2004US-00782141.

XX PR

XX 19-FEB-2004; 2004US-00782570.

XX PR

XX 19-FEB-2004; 2004US-00783417.

XX XX

XX (ATHE-) ATHENIX CORP.

XX PA

XX Carozzi N, Hargiss T, Kozziel MG, Duck NB, Carr B;

XX PI

XX WPI; 2004-635574/61.

XX XX

XX P-PSDB; ADR89397.

XX DR

XX XX

XX New isolated delta-endotoxin and delta-endotoxin-associated nucleic acids

XX PT

XX and polypeptides, useful for killing lepidopteran or coleopteran pests or

XX PT

XX for producing organisms with pesticide resistance.

XX XX

XX Claim 1; SEQ ID NO 8; 178pp; English.

XX PS

XX This sequence encodes an isolated delta-endotoxin. Some of the delta-

XX CC

XX endotoxin coding sequences of the invention have alternative start

XX CC

XX codons, producing more than one protein from a single open reading frame.

XX CC

XX The nucleic acid sequences of the invention are useful in DNA constructs

XX CC

XX or expression cassettes for transformation and expression in plants and

XX CC

XX bacteria. The nucleic acids and corresponding polypeptides are useful for

XX CC

XX killing lepidopteran or coleopteran pests. Compositions containing the

XX CC

XX delta-endotoxins of the invention, and methods for their production, are

XX CC

XX useful for the production of organisms with pesticide resistance,

XX CC

XX specifically bacteria and plants. These organisms are useful for

XX CC

XX generating altered or improved delta-endotoxin or delta-endotoxin-

XX CC

XX associated proteins that have pesticidal activity, or for detecting the

XX CC

XX presence of delta-endotoxin or delta-endotoxin-associated proteins or

XX CC

XX nucleic acids in products or organisms.

XX XX

XX SQ

XX Sequence 2235 BP; 861 A; 364 C; 316 G; 694 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 9.25e-290 Length: 2235

Score: 3314.00 Matches: 638

Percent Similarity: 91.28% Conservative: 32

Best Local Similarity: 86.92% Mismatches: 60

Query Match: 85.66% Indels: 4

DB: 13 Gaps: 3

US-10-783-417-2 (1-735) x ADR89396 (1-2235)

Qy 1 MethAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20

Db 1 GTGTAATCAAAATAAT---AATAATGAATATGAGATTATCGATTCAAGAAATTTATCTTAT 57

Qy 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40

Db 58 CTTCTTACAGAAATATTGTGATCTTCTAGATACCTTACACAAATAATCAAAATCAACCA 117

Qy 41 LeuGlnAsnThrAsnTyrLysGluThrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60

```
Db 118 TTACAAAACAAATTTACAAAGAGTGGCTCAATATGTGTCAAGGAATACACAAATGGT 177
Qy
Db 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIle 80
Db 178 GATAAATTCGAGACATTTGCTAGTGTGTATACAAATGCTGCAGTTAGTCAGGTACTATT 237
Qy 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
Db 238 GTATCCGGTACTCTGTAGCCGTATAGGTGGCTCACTTCTATATCCGACCGATAGGA 297
Qy 101 IleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPheTTPProAlaGly 120
Db 298 ATAATAGGTGCTATAATAATATCTTTTGGTACCCTTAATCACTGTCTTTTGGCCCGGGGA 357
Qy 121 GluGlnAspLysThrValTTPThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db 358 GAACAAAGACAAAACAGTATGACACAAATTTATAAATGGAGAAATTTTGTGTATACA 417
Qy 141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGlyPheArgGlnIle 160
Db 418 CCGTTAACGAAAGACATAAAACAGCTAAAGCTTACAAACTTTAGAAGGATTTAGACAAATA 477
Qy 161 LeuGlnSerTyAsnThrAlaLeuAspAspTTPArgLysLeuLysArgLeuGlnAlaPro 180
Db 478 TTACAAAGCTATAATAACAGCATTAGATGATGAGAAATTTAAAAAGACTACAAAGCTCCT 537
Qy 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGlu 200
Db 538 GGAATACCAACCATCATCAGATTACAAACAGCTGCCTTGACTCTTAAATATACGATTTGAG 597
Qy 201 AsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyLysThr 220
Db 598 AATGTTTCACAATGATTTTATTCGAGAAATACCTGGTTTCCAACTTGAAACTTATATAAAG 657
Qy 221 LeuLeuLeuProIleTyAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGly 240
Db 658 CTATTACTACCTATTATGCGCAAGCTGCTAAATTTTCATTAAATTTATTACAAACAGGT 717
Qy 241 AlaGluLeuAlaAspGluTTPAsnAlaAspIleHisProSerGlnIleGluProAsnAla 260
Db 718 GCTGAATGCGTGATGAAATGGAATGAGATACATCCCTTCAAAATTTGAACCTAATGCT 777
Qy 261 GlyThrSerAspAspTyTyLysLeuLeuLysGluAsnIleProLysTySerAsnTy 280
Db 778 GGAACATCAGATGACTATTATTAACCTTTTAAAGAAAAATATACCTAAATATAGTAATAT 837
Qy 281 CysAlaAsnThrTyArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTTP 300
Db 838 TGTGCAAAATACCTATAGAGAGGACTTAAATAAACTTCGAAACGAACTAATATGAGATGG 897
Qy 301 SerIlePheAsnAspTyArgArgTyMetThrIleThrValLeuAspThrIleSerGln 320
Db 898 AGTATATTTAATGATTATCGAAGATATAGTACTATTACTGTATTAGATACATATCGCTCAA 957
Qy 321 PheSerLeuTyAspIleLysArgTyArgAspSerIleGlyGlyIleGluValLysGly 340
Db 958 TTTTCTTTTATGATATAAGAGATACAAAGATTCATAGGA-----AGAAATAGGTGGC 1011
Qy 341 IleLysAsnGluLeuThrArgGluIleTyThrThrGluIleAsnPheAspArgLeuPro 360
Db 1012 ATTAACAACTGAACTTACAAGAGAAATTTATACAACTGAAATAAATTTTGACCGCTTACT 1071
Qy 361 GlnLeuArgValGlnProAsnLeuAlaThrMetGluTyAsnLeuThrArgAlaSerPhe 380
Db 1072 TACCTTGAATTCACCCCAATCTCGCTATAATGGAATATAATTTAACACGTTCCGGGCTT 1131
Qy 381 LysLeuPheSerPheLeuGluGlnPheIlePheTyThrThrGluAsnThrAsnPheGlyAsn 400
Db 1132 AGATTATTTTTCATTTTATAGTAGACTTATATTTTATACAAAAATGAAACGTACGGGAAT 1191
Qy 401 ArgLeuValGlyLysSerAsnArgAspAlaProThrTySerAsnThrIleThrGluThr 420
```

RESULT 3

ADR89398
ID ADR89398 standard; cDNA; 2085 BP.

XX ADR89398;
XX

DT 18-NOV-2004 (first entry)

```
Db 1192 CGTTAGTTGGTATTGCAATCGTAATAGATCTACTATGCTACGACAGGAAGTGAATTT 1251
Qy 421 LeuTyGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTy 440
Db 1252 ATATATGGAGAAAGAACAGGTCCACCCACACAAACAACTTTAATACCATTTTGAATCCTAT 1311
Qy 441 LysValSerIleValThrAspArgGlnSerProProValSerProIle--GlnProHis 459
Db 1312 AAAAGTTTCAATTTGAATGATAGACAAAGTAACTCTACTTCCCTCTTCTTAAACATATAC 1371
Qy 460 PheIleIleAsnGlnIleGluLeuTyLeuAsnGlySerSerAsnThrLeuLysTy 479
Db 1372 TTTTACAATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1431
Qy 480 SerAlaGlySerLeuSerAsnTyGlnAsnThrThrPhePheGlnPheProArgLys 499
Db 1432 TCAGCTGGGGGAATTTATCTAATGATAAAAAACAACTGATTTTCAATTTTCTTGTAATA 1491
Qy 500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTySerHis 519
Db 1492 AAAGACTGTAAACCAATTTAATCCAAATTTGTTTACCAAGCTATAATAGTTATAGTCAT 1551
Qy 520 IleLeuSerHisPheSerLeuPheThrTySerTyValIleGlyLeuGlnGlnIle 539
Db 1552 ATTTATCCAGTTTCTTTATTAAATTTCTTATAAAATTTGGATTAGCCCTAAATATA 1611
Qy 540 LeuAspThrGlyValLeuGlyTTPThrHisSerSerValAspArgTyAsnAlaIleSer 559
Db 1612 TTATATACAGTGCATTTAGGATGGACACACAGTAGTGTAAATAGAATAATGCAATATCA 1671
Qy 560 AspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLys 579
Db 1672 GATAAAATAATTAACAATGATCCAGCAATCAAGGTAACAGCTCTGTATACAACTCTAAG 1731
Qy 580 ValIleGluGlyProGlyHisThrGlyAsnLeuValTyLeuGlnSerGlnGlyArg 599
Db 1732 GTAATTTGAAGAGCTGGTCTACAGGAGAAACTTGGTTTATTATACAAAGTCAAGGGCGT 1791
Qy 600 LeuGluIleThrCysGluThrProAsnSerThrGlnSerTyPheIleArgLeuArgTy 619
Db 1792 TTAGAGATTACATGTAGAACTCTTAATCTACACAATCTTATTACATTAAGATTTCGATAC 1851
Qy 620 AlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIle 639
Db 1852 GCTACAAATGGTGTGGAATACTCTTCTTAATATATATCTTCAATACCAAGGAGTAATA 1911
Qy 640 GlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyAsnAsnLeuGln 659
Db 1912 GGAATACCCTCAACGACTCAACACACTTTTCTGGTACAAATTTATAATTTTACAA 1971
Qy 660 TyrGlyAspPheGlyTyPheGlnPheProSerThrValThrLeuProLeuAsnArgAsn 679
Db 1972 TAGGAGATTTTGGGTATTTTCCAAATTTCCAAAGTACAGTAACATTTACCTTTAAATCGA 2031
Qy 680 IleProPheIlePheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIle 699
Db 2032 ATACCATTTATATTAATCGTCAGATGTATCAAAATTTCAATTTTAAATCATGATAAAT 2091
Qy 700 GluPheIleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThr 719
Db 2092 GAATTTATACCAATTTACTCTCTGTACGCCAAATAGAGAAAAACAAAAATTGAAACT 2151
Qy 720 IleGlnThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db 2152 ATCAACAAAAATAAATACATTTTTCACAAATCATCAAAA 2193
```

XX AXMI-007 alternative start site coding sequence.

DE ss; gene; delta-endotoxin; delta-endotoxin associate polypeptide;
 KW expression cassette; transformation; transgenic; plant; bacteria;
 KW lepidoptera; coleoptera; pest; pesticide; resistance;
 KW pesticidal activity.

XX Bacillus thuringiensis.

XX Key Location/Qualifiers

FT CDS 1. 2085

FT /*tag= a

FT /product= "Alternative AXMI-007"

XX WO2004074462-A2.

XX 02-SEP-2004.

XX 20-FEB-2004; 2004WO-US005829.

XX 20-FEB-2003; 2003US-0448632P.

XX 20-FEB-2003; 2003US-0448633P.

XX 20-FEB-2003; 2003US-0448797P.

XX 20-FEB-2003; 2003US-0448806P.

XX 20-FEB-2003; 2003US-0448810P.

XX 20-FEB-2003; 2003US-0448812P.

XX 19-FEB-2004; 2004US-00781979.

XX 19-FEB-2004; 2004US-00782096.

XX 19-FEB-2004; 2004US-00782141.

XX 19-FEB-2004; 2004US-00782570.

XX 19-FEB-2004; 2004US-00783417.

XX (ATHE-) ATHENIX CORP.

XX Carozzi N, Hargiss T, Kozziel MG, Duck NB, Carr B;
 XX WPI; 2004-635574/61.

XX P-PSDB; ADR89399.

XX New isolated delta-endotoxin and delta-endotoxin-associated nucleic acids
 PT and polypeptides, useful for killing lepidopteran or coleopteran pests or
 PT for producing organisms with pesticide resistance.

XX Claim 1; SEQ ID NO 10; 178pp; English.

XX This sequence encodes an isolated delta-endotoxin. Some of the delta-
 CC endotoxin coding sequences of the invention have alternative start
 CC codons, producing more than one protein from a single open reading frame.
 CC The nucleic acid sequences of the invention are useful in DNA constructs
 CC or expression cassettes for transformation and expression in plants and
 CC bacteria. The nucleic acids and corresponding polypeptides are useful for
 CC killing lepidopteran or coleopteran pests. Compositions containing the
 CC delta-endotoxins of the invention, and methods for their production, are
 CC useful for the production of organisms with pesticide resistance,
 CC specifically bacteria and plants. These organisms are useful for
 CC generating altered or improved delta-endotoxin or delta-endotoxin-
 CC associated proteins that have pesticidal activity, or for detecting the
 CC presence of delta-endotoxin or delta-endotoxin-associated proteins or
 CC nucleic acids in products or organisms.

XX Sequence 2085 BP; 794 A; 338 C; 302 G; 651 T; 0 U; 0 Other;

XX Alignment Scores:

Pred. No.:	7.9e-271	Length:	2085
Score:	3104.50	Matches:	598
Percent Similarity:	91.95%	Conservative:	30
Best Local Similarity:	87.55%	Mismatches:	52
Query Match:	80.24%	Indels:	3
DB:	13	Gaps:	2

XX US-10-783-417-2 (1-735) x ADR89399 (1-2085)

Qy	52	MetCysGlnGlyAsnThrGlnTyrGlyAspAsnPheGluThrPheAlaSerAlaSepThr	71
Db	1	ATGTGTCAAGGGAATACACAATATGGTGAATATTCGAGACATTTGCTAGTCTGATACA	60
Qy	72	IleAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAlaGlyIleGlyGly	91
Db	61	ATTGCTGCAGTTAGTGCAGGTACTATTGTATCCGGTACTCTGTTAGCCGGTATAGTGG	120
Qy	92	LeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIleIleSerPheGlyThr	111
Db	121	CTCACTTCTATATCCGGACCGATAGGAATAATAGTGCTATAATAATAATCTTTTGTACC	180
Qy	112	LeuIleThrValPheTyrProAlaGlyGlnAspLysThrValTyrThrGlnPheIle	131
Db	181	CTAATCACTGTCTTTTGGCCCGCGGAGAAACAAGACAAAGATATGGACACAATTTATT	240
Qy	132	LysMetGlyGluIlePheValAspThrProLeuThrGluSerIleLysGlnLeuLysLeu	151
Db	241	AAATGGGAGAAATTTTGTGATACACCGTTACAGAAAGCATAAACAGCTAAAGTTA	300
Qy	152	GlnThrLeuGluGlyPheArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspThr	171
Db	301	CAAACTTTAGAGGATTAGACAAATATTACAAGCTATAATACAGCATTAGATGATGG	360
Qy	172	ArgLysLeuLysArgLeuGlnAlaProGlyLeuProProSerSerAlaLeuGlnAla	191
Db	361	AGAAAATTTAAAGACTACAAGCTCTTGGATTTACCACCATTATCAACAAGCT	420
Qy	192	AlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPheIleArgGluIlePro	211
Db	421	GCCTTGACTCTTAAATACGATTTGAGAAATGTTTCAATGATTTTATTCGAGAAATACCT	480
Qy	212	GlyPheGlnLeuGluThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAsn	231
Db	481	GGTTTCCAACTTGAACCTTATAAAGCTATTACTACCTATTATTCGCAAGCTGCTAAT	540
Qy	232	PheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTyrAsnAlaAspIle	251
Db	541	TTTCATTTAAATTTATTACAACAAGGTCTGAAATGGCTGATGAATGGAATGCAGATATA	600
Qy	252	HisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTyrTyrLysLeuLys	271
Db	601	CATCCTTCAACAATTTGAACCTTAATGCTGGAACATCAGATGACTATTATTAACCTTTAAA	660
Qy	272	GluAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsn	291
Db	661	GAAATATACCTAATATAGTAATTTGTGCAAAATACCTATAGAGAGGACTAATAATA	720
Qy	292	LeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTyrArgTyrMetThr	311
Db	721	CTTCGAAACGAACCTAATATGAGATGGAGTATATTATTAATGATTCGAAGATATATGACT	780
Qy	312	IleThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAsp	331
Db	781	ATTAAGTGTATTAGATACCTATCGCTCAATTTCTTTTATGATATATAAGAGATACAAAGAT	840
Qy	332	SerIleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThr	351
Db	841	TCAATAGA-----AGATAGGTGGCATTAATACTGAAGCTTACAGAGAAATTTATACA	894
Qy	352	ThrGluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMet	371
Db	895	ACTGAATAAATTTTGACCGTCTTACTTTACCTTGAATTCACCAATTCGCTATAATG	954
Qy	372	GluTyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePhe	391
Db	955	GAATATAAATTTAACACGTTACAGGCTTACAGGCTTACAGGCTTACAGGCTTACAGGCTT	1014
Qy	392	TyrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIleSerAsnArgAspAlaPro	411
Db	1015	TATACAAAAATGAACGTCACGGGAATCGTTTAGTTGGTATTGCGAATCGTAATAGATCT	1074

Db 1065 GGAGATTTTGAACCTTTTATTGATAGT-----GGTGAACCTCAGTCCCTATACTATT 1115
Qy 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
Db 1116 GTAGTTGGACCGACTGACTGGTTTCGGGTTCAACA-----CCCTTAGGA 1163
Qy 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120
Db 1164 CTT-----GCTTTAATAGGTTTGGTACATTAATACAGTCTCTTTCCAGCCCA 1214
Qy 121 GluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db 1215 GACCAATCAACACA-----TGGAGTGACTTTATACACAACTAAATAATATTATAAAAA 1271
Qy 141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
Db 1272 GAAATGACATCAACATATATAAGTAATGCTAATAAAATTTTAAACAGGTCGTTTAATGTT 1331
Qy 161 LeuGlnSerTrpAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaPro 180
Db 1332 ATCAGCACTTATCATATACCTTAAACATGGAG-----AATAATCCA 1376
Qy 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGlu 200
Db 1377 AACCCACAAATACTCAGGATGTAAGGACACAAATCCAGCTAGTTCATTACCAITTTCAA 1436
Qy 201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTrp 218
Db 1437 AATGTCAATCCAGAGCTTTAAACTCTTGCTCCTAATCTAGTGTGATTCGATTAAT 1496
Qy 219 LysThrLeuLeuLeuProIleIleTyAlaGlnAlaAlaAsnPheHisLeuAsnLeuGln 238
Db 1497 AACATACGTATTAATCTAGTTATGCAACAGCAGCAAACTTACATCTGACTGTATTAAT 1556
Qy 239 GlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluPro 258
Db 1557 CAAGCCGTCAAATTTGAAGCGTATTTAAACAAATCGACAAATCGATTATTAGAGCT 1616
Qy 259 AsnAlaGlyThrSerAspAspTyTrpLysLeuLeuLysGluAsnIleProLysTrpSer 278
Db 1617 ---TTGCCAACAGCAATTTGATTATTAATCCAGTATTCAGTAAAGCTATAGAGATTACACT 1673
Qy 279 AsnTyrcysAlaAsnThrTrpArgThrGlyLeuLysAsnLeuArgAspGluPro----- 296
Db 1674 AATTATTGTGTAAACAACCTTAAACAAAGGATTAATTTAATTAACAGCAGCCTCATAGT 1733
Qy 297 -----AsnMetLysTrpSerIlePheAsnAspTrpArgArgTrpMetThrIle 312
Db 1734 AATCTTGATGGAATATAAAGCTGGAACACATACATACATAGTATCGAACAAATGACTACT 1793
Qy 313 ThrValLeuAspThrIleSerGlnPheSerLeuTrpAspIleLysArgTrpArgAspSer 332
Db 1794 GCTGTATTAGATGTTGTGGACTCTTTCTTAATTAATGATGATGATGATGATGATGAT 1844
Qy 333 IleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyThrThr 352
Db 1845 -----CCAATAGGTGTCCAAATCTCAACTTACTCAGAGAAATTTATCAGGTA 1899
Qy 353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
Db 1890 ---CTTAACCTTCGAAGAAAGCCCTATAAATATTAT-----GACTTCATATATCAAGAG 1940
Qy 373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTy 392
Db 1941 GATTCACCTTACCGTAGACCG---CATTATTACTTGGCTTGATCTCTTGAATTTTAT 1997
Qy 393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
Db 1998 GAAAAAGCGCAAACTACTCTCTAATAATTTTTCACGAGCCATTAATATATGTTTCATTAC 2057
Qy 402 ---LeuValGlyIleSerAsnArgAspAlaProThrTrpSerAsnThrIleThrGluThr 420
Db 2058 ACACCTTGATAATATATCCAAAAATCTAGTGTGTTTTTGGAAATCAATGTAACCTGATAAA 2117

RESULT 5

AAQ14669

ID AAQ14669

standard; DNA; 3543 BP.

XX

Qy 421 LeuTyrglyGluArgThrGlySerProThrThrThrLysThrIleArgProPheGluSerTyr 440
Db 2118 TTA-----AAATCTCTTGGTTTGGCAACAATAATTATATATT----- 2153
Qy 441 LysValSerIleValThrAspArgGlnSerProValSerProIleGlnProHisPhe 460
Db 2154 -----TTT 2156
Qy 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
Db 2157 TTATTAAATGTCATAAAGCTTAGATAATAAATATCTAAATGATTATAATAATTAGTAAA 2216
Qy 477 -----LeuLysTyrSerAlaGlyGly 483
Db 2217 ATGGATTTTTTATAACTAATGTTACTAGACTTTTGGAGAAAGAACTTACAGCAGGATCT 2276
Qy 484 SerLeuSerAsnTyrgln---AsnThrThrPhePheGlnPheProArgLysLysAspCys 502
Db 2277 GGGCAATAACTATTATGATGTAATAATAAATAATTTTCGGGTTACCAATCTTTAAACGAAGA 2336
Qy 503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrsSerHisIleLeuSer 522
Db 2337 GAGAAATCAAGGAACCCCTACCCCTTTTCCACATATGATACTATAGTTCATATTTATCA 2396
Qy 523 HisPheSerLeuPheThrTrpSerTyrvallleGlyLeuGlnLeuIleLeuAspThr 542
Db 2397 TTTATTAAAGTCTTAGTATCCCTGCAACATATAAACTCAAGTGTAT----- 2444
Qy 543 GlyValLeuGlyTrpThrHisSerSerValAspArgTyrsAsnAlaIleSerAspLysIle 562
Db 2445 ---ACGTTTGTGTGACACACTCTAGTGTGTATCTATAAAATACAAATTTATACACATTTA 2501
Qy 563 IleThrMetIleProAlaIleLysGlyAsnLeuAspThrAsnSerLysValIleGlu 582
Db 2502 ACTACCCAAATTCAGCTGTAAAGCGAATTCACCTTGGGACTGCTCTTAAGGTGTGTCAA 2561
Qy 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrglyLeuGlnSerGlnGlyArgLeuGluIle 602
Db 2562 GGAGCTGTCTATACAGGAGGGAATTAAT-----GATTTCAAAAGATCATTTCAAAAT 2615
Qy 603 ThrCysGluThrProAsnSerThrGlnSerTyrglyPheIleArgLeuArgTyrglyAlaThrAsn 622
Db 2616 ACATGTCAACACTCAAAATCTCAACAAATCGTATTTTATAAGAAATTCGTTATGCTTCAAAT 2675
Qy 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db 2676 GGAAGCGCAATATCTCAGAGCTGTATAAATCTTAGTATCCAGGGGTAGCAGAACTG--- 2732
Qy 643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrglyAsnLeuGlnTyrglyAsp 662
Db 2733 GGTATGGCACTCAACCCCACTTTTTCGTGTCAGATTTATACGAATTTAAATAATAAAGAT 2792
Qy 663 PheGlyTyrglyPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
Db 2793 TTTCAGTACTTAGAATTTTCTAACGAGGTGAAATTTGCTCCAAATCAAAACATATCTCT 2852
Qy 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleAspLysIleGluPhe 701
Db 2853 GTGTTTAATCTGCGATGATATACAAACACACACAGTACTTATTGATATAAATTTGAATTT 2912
Qy 702 IleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGln 721
Db 2913 CTGCCAATTAATCTGTTCTTAAGAGAGGATAGAGAGAAACAAATAATAGAACAGTACAA 2972
Qy 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db 2973 CAATAAATTAATACATTTTATGCAAAATCCTATAAAA 3008


```
Db 274 CTT-----GCTTTAATAGCTTTTGGTACATTAATACCAGTCTCTTTTTCAGGCCAA 324
Qy 121 GluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluPheValAspThr 140
Db 325 GACCAATCTAACACA---TGGAGTGACTTTATACACAACTAAATATATATAAAAAA 381
Qy 141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
Db 382 GAAATAGCATCAACATATATAAGTAATGCTAATAAAAAATTTTAAACAGGTGCTTAAATGTT 441
Qy 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaPro 180
Db 442 ATCAGCACTTATCATAAATCACCTTAAACATGGGAG-----AATAATCCA 486
Qy 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGlu 200
Db 487 AACCCACAAAATACTCAGGATGTAAGGACACAAATCCAGCTAGTTCATTACCAATTTTCAA 546
Qy 201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr 218
Db 547 AATGTCATTCAGAGCTTGTAACCTCTTGCTCCTTAATCCTAGTATGCGATTACTAT 606
Qy 219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGln 238
Db 607 AACATACTAGTATTATCTAGTTATGCAAGCAGCAAACTTACATCTGACTGATTAAAT 666
Qy 239 GlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluPro 258
Db 667 CAAGCGCTCAAAATTTGAAGCGTATTTAAAAACAATCGACAATTCGATTATTTAGAGCCT 726
Qy 259 AsnAlaGlyThrSerAspAspTyrTyrLysLeuLysGluAsnIleProLysTyrSer 278
Db 727 ---TTGCCAACAGCAATTCGATTATATCCAGTATTCATTAAGCTATAGAGATTACACT 783
Qy 279 AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuAspGluPro----- 296
Db 784 AATTATGTGTACAACTTATAAAAAAGGATTAAATTTAATTAACAGCGCTGTAGT 843
Qy 297 -----AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIle 312
Db 844 AATCTTGATGGAATATAAATCGAACACACATCAATCAATCGTATCGAACAATAATGACTACT 903
Qy 313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer 332
Db 904 GCTGATTAGATCTGTTGCACCTCTTCCTAATATATGATGATGATTAAT----- 954
Qy 333 IleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThr 352
Db 955 -----CCAATAGGTGTCCTCAATCTGAACCTTACTCGAGAAATTTATCAGGTA 999
Qy 353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
Db 1000 ---CTTAACTTCGAAGAAAGCCCTATAAATATAT-----GACTTTCATATCAAGTA 1050
Qy 373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyr 392
Db 1051 GATTCACTTACAGTACAGCG---CAATTATTTACTTGGCTTGATCTTGGAAATTTTTAT 1107
Qy 393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
Db 1108 GAAAGGCGCAAACTACTCTCTAATAATTTTTCACAGCCATTAATAATATGTTTCATTAC 1167
Qy 402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
Db 1168 ACATTGATAATATATCCCAAAATCTAGTGTGTTTGGAAATCACAATGTAACCTGATAAA 1227
Qy 421 LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440
Db 1228 TTA-----AAATCTCTGTTGGCAACAAATATTTATATTT----- 1263
Qy 441 LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
```

```
Db 1264 ----- 1266
Qy 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
Db 1267 TTATTAATGTCATAAGCTTAGATAATAAATATCTAATGATTATAATAATATTAGTAAA 1326
Qy 477 -----LeuLysTyrSerAlaGlyGly 483
Db 1327 ATGGAATTTTTTATAAATAATGTTAGTACTTTTGGAGAAAGAACTTACAGCAGATCT 1386
Qy 484 SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysLysAspCys 502
Db 1387 GGGCAATAACTATTATGATGTAATAAATAAAYATTTTCGGGTACCAATCTTAAACCAAGA 1446
Qy 503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
Db 1447 GAGAAATCAAGCAATCCCTACCCCTTTTCCACATATGATATATGTCATATTTTATCA 1506
Qy 523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThr 542
Db 1507 TTTATTAATAAGTCTTAGTATCCCTGCAACATATAAAACTCAAGTGAT----- 1554
Qy 543 GlyValLeuGlyTrpThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
Db 1555 ---ACGTTTGGTGGACACACTCTAGTGTGTGATCCTTAAAAATACAAATTTATACACATTTA 1611
Qy 563 IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582
Db 1612 ACTACCAAAATTCAGCTGTAAAGCGAATTCACCTGGGACTGCTTCTAGGTTGTTCAA 1671
Qy 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle 602
Db 1672 GGACCTGTCATACAGGAGGGGATTTAAT-----GATTTCAAAGATCATTTCAAAAT 1725
Qy 603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
Db 1726 ACATGTCAACACTCAAAATTTTCAACAATCGTATTTATTAAGAAATTCGTATGCTTCAAA 1785
Qy 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db 1786 GGAAGCCAAATACACGAGCTGTATTAATCTTAGTATCCAGGGGTAGCAGAACTG--- 1842
Qy 643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
Db 1843 GGTATGCACTCAACCCACCTTTTCTGTCACAGATTATACGAATTTAAATATAAAGAT 1902
Qy 663 PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
Db 1903 TTTCACTACTTAGAATTTTCTAACGAGGTGAAATTTGCTCCAAATCAAAACATATCTCTT 1962
Qy 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleAspLysIleGluPhe 701
Db 1963 GTGTTTAACTGTCGGATGATATATACAAACACACAGTACTTATTGATAAAATTTGAATTT 2022
Qy 702 IleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGln 721
Db 2023 CTGCCAATTAATCTGTTCTATTAAGAGAGATAGAGAGAAACAAATTTAGAACAGTACAA 2082
Qy 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db 2083 CAATAATAATAATATATTTATGCAAAATCCTATAAAA 2118
RESULT 7
AAN93054
ID AAN93054 standard; DNA; 3940 BP.
XX
AC AAN93054;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 30-MAY-1990 (first entry)
XX
DE Delta-endotoxin crystal protein gene.
```

XX Delta-endotoxin; crystal protein; insecticide; ss; pCC130;
 KW biological control agent.
 XX
 OS *Bacillus thuringiensis*; israelensis.
 FH Location/Qualifiers
 FT Key 879..884
 FT RBS /*tag= b
 FT /label= Shine-Delgarno sequence
 FT CDS 891
 FT /*tag= a
 FT /product= "delta-endotoxin"
 XX
 PN EP296970-A.
 XX
 PD 28-DEC-1988.
 XX
 PF 24-JUN-1988; 88EP-00305772.
 XX
 PR 26-JUN-1987; 87US-00067653.
 XX
 XX (DUPO) DU FONT DE NEMOURS & CO E I.
 PA
 XX Ellar DJ, Ward ES;
 PI
 XX WPI; 1989-001322/01.
 DR P-PSDB; AAP94035.
 XX
 XX DNA fragment encoding insecticidal protein - obtd. from *Bacillus*
 PT *thuringiensis* sub species *israelensis*, and used in microorganisms and
 PT plant cells.
 XX
 XX Disclosure; Fig 5; 26pp; English.
 XX
 CC The nucleotide sequence is an insert in plasmid pCC130. The delta
 CC endotoxin protein is insecticidal and can be used to control insect pests
 CC esp. mosquitoes. See also AAN93059. (Updated on 25-MAR-2003 to correct PA
 CC field.) (Updated on 27-AUG-2003 to correct OS field.)
 XX
 SQ Sequence 3940 BP; 1457 A; 603 C; 659 G; 1221 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 3.06e-85 Length: 3940
 Score: 1059.50 Matches: 271
 Percent Similarity: 50.26% Conservative: 117
 Best Local Similarity: 35.10% Mismatches: 279
 Query Match: 27.38% Indels: 105
 DB: 1 Gaps: 24
 US-10-783-417-2 (1-735) x AAN93054 (1-3940)
 QY 1 MetAsnGlnAsnAspAsnAsnGluTyrGluilelleAspSerHisThrSerProTyr 20
 DB 891 ATGAATCCTTATCAAAATTAATAATGAATATGAATGAACATTTAAATGCTTCACAAAAAAATTA 950
 QY 21 PheProAsnArgAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
 DB 951 -----AATATATCTAATAATATATACAGATATCCATAGAAATAGTCCAAAAACAATTA 1004
 QY 41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
 DB 1005 TTACAAAGTACAAATTAATAAAGATTGGCTCAATATGTGTCAACAGATCAGCAGTATGGT 1064
 QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIle 80
 DB 1065 GGAGATTTTGAACCTTTTATTTAGTAGT-----GGTGAACCTCAGTGCCTATACTATT 1115
 QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
 DB 1116 GTAGTTGGGACCGTACTGCTGGTTTCGGGTTCAACA-----CCCTTAGGA 1163
 QY 101 IleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGly 120

DB 1164 CTT-----GCTTTAATAGTTTGGTACATTAATACCGATCTTTTTCAGCCCAA 1214
 QY 121 GluGlnAspLysThrValThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
 DB 1215 GACCAATCTAACACA---TGGAGTGACTTTATAACACAACTAAAAATATTATAAAAAA 1271
 QY 141 ProIleThrGluSerIleLysGlnIleLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
 DB 1272 GAAATAGCATCAACATATATAAGTAATGCTAATAAAATTTTAAACAGGTCGTTTAATGTT 1331
 QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTTPArgLysLeuLysArgLeuGlnAlaPro 180
 DB 1332 ATCAGCACTTATCATATAATCACCTTAAACATGGGAG-----AATAATCCA 1376
 QY 181 GlyLeuProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGlu 200
 DB 1377 AACCCACAAAATACTCAGGATGTAAGGACACAAATCCAGCTAGTTTCATTCACCAATTTCAA 1436
 QY 201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr 218
 DB 1437 AATGTCAATCCACAGCTTGTAACTCTTGTCTCTCTAATCCTAGTAGTTCGATTCGATTA 1496
 QY 219 LysThrLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGln 238
 DB 1497 AACATACTAGTATTATCTAGTTATGCACAGCAGCAACTTACATCTGACTGTATTAAAT 1556
 QY 239 GlnGlyAlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluPro 258
 DB 1557 CAAGCGCTCAAAATTTGAAGCGTATTTAAAAACAATCGACAATTCGATTATTAGAGCCT 1616
 QY 259 AsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSer 278
 DB 1617 --TTGCCACACAGCAATTTGATTATATCCAGTATTGACTTAAAGCTATAGAAGATTACACT 1673
 QY 279 AsnTyrCysAlaAsnThrTyrArgThrGlyIleLysAsnLeuArgAspGluPro----- 296
 DB 1674 AATTATTGTGTAAACAACCTTATAAAAGGATTAAATTTAATTTAAACACGCGCTGATAGT 1733
 QY 297 -----AsnMetLysTyrSerIlePheAsnAspTyrArgArgTyrMetThrIle 312
 DB 1734 AATCTTGATGGAAATATATAAATCTGGAACATACATACATACGTATCGAACAATAATGACTACT 1793
 QY 313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer 332
 DB 1794 GCTGTATTAGATGTTGTTGCACTCTTCTCTAATTATGATGAGTAAATAT----- 1844
 QY 333 IleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThr 352
 DB 1845 -----CCAAATAGGTGTCCAATCTGAACCTTACTCGAATAATTTATCAGGTA 1889
 QY 353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
 DB 1890 ---CTTAATCTCGAAGAAAGCCCTATATAATATTAT-----GACTTTCAATATCAAGAG 1940
 QY 373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnPheIlePheTyr 392
 DB 1941 GATTCACCTTACCGTAGACCG---CATTATTATTACTTGGCTTGATCTTTGAATTTTAT 1997
 QY 393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
 DB 1998 GAAAAAGCGCAAACTACTCTCTAAATAATTTTTCACGAGCCATATAATATGTTTCATTAC 2057
 QY 402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
 DB 2058 ACACCTTGATAATATATCCAAAAATCTAGTGTGTTTGGAAATACCAATGTACTATGATAA 2117
 QY 421 LeuTyrGlyGluArgThrGlySerProThrLysThrIleArgProPheGluSerTyr 440
 DB 2118 TTA-----AAATCTCTGGTTTGGCAACAATAATTTATATT----- 2153
 QY 441 LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460

PN JP2004166574-A.
 XX 17-JUN-2004.
 XX 19-NOV-2002; 2002JP-00335440.
 XX 19-NOV-2002; 2002JP-00335440.
 PR (UYOK-) UNIV OKAYAMA.
 PA
 XX WPI; 2004-445549/42.
 XX
 PT Novel codon-modified gene encoding Cry4A protein, the modification being
 PT substitution with most frequently used codons in host microorganisms such
 PT that resulting gene has high translation rate in the microorganism.
 XX
 XX Claim 3; SEQ ID NO 1; 20pp; Japanese.
 XX
 CC The invention comprises a codon-modified gene which codes for Cry4A
 CC protein, the gene is modified by substitution with the most frequently
 CC used codon in the host microorganism, so that the modified gene will
 CC a high translation rate in the host microorganism. The codon-modified
 CC gene of the invention is useful for the production of Cry4A protein in
 CC large quantities. The present DNA sequence represents a codon-modified
 CC Cry4A gene of the invention.
 XX
 SQ Sequence 2130 BP; 677 A; 704 C; 428 G; 321 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3,64e-83 Length: 2130
 Score: 1032.50 Matches: 265
 Percent Similarity: 50.26% Conservative: 117
 Best Local Similarity: 34.87% Mismatches: 273
 Query Match: 26.69% Indels: 105
 DB: 12 Gaps: 24

US-10-783-417-2 (1-735) x ADP71294 (1-2130)

Qy 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLeuLeuLeuLeuLeu 200
 Db 516 AACCGCAAAACACCCCAAGCGTGGTACCCAAATCCAACTGGTGCACTACCACTTCCAA 575
 Qy 201 AsnValHisAsnAspPheIleArgGluLeuPro-----GlyPheGlnLeuGluThrTyr 218
 Db 576 AACGTTCATCCCGGAACCTGGTGAACAGCTGCCCGCGAAACCCGAGCGACTGGCGACTAC 635
 Qy 219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGln 238
 Db 636 AACATCTCGTGTGACGAGCTAGCGCGAGCAGCGAACCTGCACTGACCGTGTGAAC 695
 Qy 239 GlnGlyAlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluPro 258
 Db 696 CAACGGGTCAAAATTCGAGGCTTACCTGAAACCAACCAACCAATTCGACTACCTCGAGCG 755
 Qy 259 AsnAlaGlyThrSerAspAspTyrTyrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 278
 Db 756 ---CTGCCGACCGCAATCGACTACTACCGGTGCTGACCAAGCAATCGAAGACTACACC 812
 Qy 279 AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLeuLeuLeuLeuLeuLeuLeuLeu 296
 Db 813 AACTACTGCTGACCACTACCAAAAGGCTGAACTGATCAAAACCAACCCCGGACGACG 872
 Qy 297 -----AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIle 312
 Db 873 AACTGACGCGCAACATCACTGAAACACCTACACCACTACCGCAACCAAAATGACCA 932
 Qy 313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer 332
 Db 933 GCGGTGCTGACCTGCTGGCACTGTTCCGAACCTACGACGTCGGTAATAATAC----- 983
 Qy 333 IleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluLeuIleTyrThr 352
 Db 984 -----CCGATCGGTGTCAAAGCAACTGACCCGGGAATCTACCAAGTC 1028
 Qy 353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
 Db 1029 ---CTGAACCTCGAAGAACCCGTCACAAATACTAC-----GACTTCCAATACCAAGAA 1079
 Qy 373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyr 392
 Db 1080 GACAGCTGACCGCTGCGCCG---CACGTGTTCACCTGGCTGGACAGCGCTGAATCTCTAC 1136
 Qy 393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
 Db 1137 GAAAAAGCGCAAAACCCCGCAACAACTTCTTCCAGCCCACTACAAATGTTCCACTAC 1196
 Qy 402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
 Db 1197 ACCCTGACCAACATCAGCCAAACCAAGCGGTGTCGCAACCAACACGTCGACGCAAA 1256
 Qy 421 LeuTyrGlyGluArgThrGlySerProThrThrIleThrIleArgProPheGluSerTyr 440
 Db 1257 CTG-----AAAAGCTGGGTCTGCAACCAACCACTATCATC----- 1292
 Qy 441 LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
 Db 1293 ----- 1295
 Qy 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
 Db 1296 CTGCTGAACGTCATCAGCTCGACCAACAAATACCTGAAAGCACTACCAACATCAGCAAA 1355
 Qy 477 -----LeuLysTyrSerAlaGlyGly 483
 Db 1356 ATGGACTTCTTCATCAACCAACGCTACCGCTGCTGCAAAAGAACTGACCGGAGCAGC 1415
 Qy 484 SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysLysAspCys 502
 Db 1416 GSCCAAAATCACCTACGACGTGAACAAACAAATCTTTCGCGCTGCCGATCCTGAACGTCGC 1475

```
QY 503 AsnLeuValIleAepProGlyCysSerProAenPheAenAenTyrSerHisIleLeuSer 522
Db 1476 GAAACCAAGCAGCCGACCTGTTCCGACCTACGACACTACGACCATCTGAGC 1535
QY 523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuLeuAspThr 542
Db 1536 TTCATCAAAAGCCTGAGCATCCCGCAACCTACAAACCCCAAGGTAC----- 1583
QY 543 GlyValLeuGlyThrHisSerSerValAspArgTyrAenAlaIleSerAspLysIle 562
Db 1584 ---ACCTTCGCGTGGAGCCACAGCAGCGTGCACCCGAAACACCATCTACACCCACTG 1640
QY 563 IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAenSerLysValIleGlu 582
Db 1641 ACCACCAATCCCGCGGTGAAGCAACAGCTGGCGCAGCGAGCAAGTGTGCA 1700
QY 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle 602
Db 1701 GGTCCGGGTACACCGCGCGGTGACCTGATC-----GACTTCAAGAGCCACTTCAAAATC 1754
QY 603 ThrCysGluThrProAenSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
Db 1755 ACCTGCCAACACAGCAACTTCCCAACAAAGCTACTTCATCGCATCCGTTACGCGAGCAAC 1814
QY 623 GlyValGlyAsnThrLeuProAenIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db 1815 GGACGCCAACACCCCGCGGTGATCAACTGAGCATCCCTGGCGTGGCAGAACTG--- 1871
QY 643 ProGlnArgLeuAenAsnThrPheSerGlyThrAsnTyrAenAsnLeuGlnTyrGlyAsp 662
Db 1872 GGTATGCACTGAACCGGACCTTCAGCGGTACCGACTACCAACCTGGAATACAAAGAC 1931
QY 663 PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAenArgAsnIleProPhe 682
Db 1932 TTCATATACCTGGAGTTTCAGCAACGAAGTGAAATTCGCGCGAACCACAAACATCAGCGCTG 1991
QY 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleAspLysIleGluPhe 701
Db 1992 GTTTCACCGTACGACGCTGTACCAACACACACCGTGTGATCGCAAAATTCGAGTTC 2051
QY 702 IleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGln 721
Db 2052 CTGCGGATCACCGTAGCATCCGCGAAGACCGTGAATAACAAACAACTGGAACCGTGCA 2111
RESULT 10
ADP71295
ID ADP71295 standard; DNA; 2246 BP.
XX
AC ADP71295;
XX
DT 09-SBP-2004 (first entry)
XX
XX Codon-modified Cry4A gene #2.
XX
DE codon optimisation; codon-modified; Cry4A; gene; ds.
XX
XX Synthetic.
OS Unidentified.
XX
XX JP2004166574-A.
FN
XX
PD 17-JUN-2004.
XX
XX 19-NOV-2002; 2002JP-00335440.
XX
PR 19-NOV-2002; 2002JP-00335440.
XX
XX (UYOK-) UNIV OKAYAMA.
XX
XX WPI; 2004-445549/42.
XX
XX Novel codon-modified gene encoding Cry4A protein, the modification being
PT substitution with most frequently used codons in host microorganisms such
```

```
PT that resulting gene has high translation rate in the microorganism.
XX
PS Disclosure; SEQ ID NO 2; 20pp; Japanese.
XX
CC The invention comprises a codon-modified gene which codes for Cry4A
CC protein, the gene is modified by substitution with the most frequently
CC used codon in the host microorganism, so that the modified gene will have
CC a high translation rate in the host microorganism. The codon-modified
CC gene of the invention is useful for the production of Cry4A protein in
CC large quantities. The present DNA sequence represents a codon-modified
CC Cry4A gene of the invention.
XX
```

Sequence 2246 BP; 711 A; 728 C; 457 G; 350 T; 0 U; 0 Other;

Alignment Scores:

Pred. NO.:	3.92e-83	Length:	2246
Score:	1032.50	Matches:	265
Percent Similarity:	50.26%	Conservative:	117
Best Local Similarity:	34.87%	Mismatches:	273
Query Match:	26.69%	Indels:	105
DB:	12	Gaps:	24

US-10-783-417-2 (1-735) x ADP71295 (1-2246)

```
QY 1 MetAenGlnAenAenAenAenAenGluTyrGluIleIleAenSerHisThrSerProTyr 20
Db 146 ATGAACCGGTACCAAAACAAAGACGAATACGAAACCCCTGGAACGCGAGCAAGAACTG 205
QY 21 PheProAenArgAenSerAenAenAenSerArgTyrProTyrThrAsnAenProAenGlnPro 40
Db 206 -----AACATCAGCAACACTACACCGGTACCGATCGAANAACAGCCGCAACAACTG 259
QY 41 LeuGlnAenThrAenTyrLysLeuTrpLeuAenMetCysGlnGlyAenThrGlnTyrGly 60
Db 260 CTGCAAGACCAACTACAAAGACTGGCTGAACATGTGCCAACAAACCAACAATACGCGC 319
QY 61 AspAenPheGluThrPheAenAenAenAenAenAenAenAenAenAenAenAenAenAen 80
Db 320 GCGCACTTCGAACCTTCATCGACGAC-----GGTGAACCTGAGCGGTACACCATC 370
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGly 100
Db 371 GTGTGCGCACCGGTGCTGACCGGTTCGCTTCACCACTC-----CGCTGGGC 418
QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpAlaGly 120
Db 419 CTG-----GCGCTGATCGGTTTCGTACCTCGTATCCCGTCTGCTTCCCGGCGCA 469
QY 121 GluGlnAspLysThrValThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db 470 GACCAAGCAACACG---TGGAGCGACTTCATACCCCAACCCCAAAACATCATCAAAAAA 526
QY 141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
Db 527 GAAATCGCAAGCACTTACATCAGCAACGCGCAACAAATCTCGAACCCTGAGTTCACCGTG 586
QY 161 LeuGlnSerTyrAenThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaPro 180
Db 587 ATCAGCACCTACCAACACCACTTGAACAACTGGGA-----AACACCCG 631
QY 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGlu 200
Db 632 AACCCGCAAAACACCCCAAGACGTGCGTACCCCAATCCAACTGGTGCACCTTCCAA 691
QY 201 AsnValHisAenAenPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr 218
Db 692 AACGTCATCCCGAAGCTGGTGAACAGCTGCGCGCAACCCCGAGGAGCTGCGACTACTAC 751
QY 219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAenPheHisAenLeuLeuGln 238
Db 752 AACATCTGGTGGTGGCAGCTACCGCAAGCAGCAACTCGCACTGACCGTGGCTGAAC 811
QY 239 GlnGlyAlaGluLeuAlaAspGluTrpAenAlaAspIleHisProSerGlnIleGluPro 258
```

```
Db      812  CAAAGCGTCAATTCGAGGCTTACCTGAAACCAACCGCCAAATTCGACTACCTCGAGCCG 871
Qy      259  AsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSer 278
Db      872  ---CTGCCGACCGCAATCGACTACTACCGGTGCTGACCAAGCAATCGAAGACTACACC 928
Qy      279  AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro----- 296
Db      929  AACTACTGGGTGACCACTACAAAAAGGCTGAACCTGATCAAAACCAACCCCGGACAGC 988
Qy      297  -----AsnMetLysTyrTrpSerIlePheAsnAspTyrArgArgTyrMetThrIle 312
Db      989  AACCTGGACGGCAACATCAACTGGACACACTCAACACACTTACCGCACCAAAATGACCA 1048
Qy      313  ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer 332
Db      1049  CGGTGCTGGACCTGGTGGACGTGTTCCGAACCTACGACGTGGTAATAC----- 1099
Qy      333  IleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThr 352
Db      1100  -----CGATCGGTGTCCAAAGCGAACTGACCCGGGAAATCTACCAAGTC 1144
Qy      353  GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
Db      1145  ---CTGAACCTCGAAGAAAGCGGTACAAATAC-----GACTTCCAAATACCAAGAA 1195
Qy      373  TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyr 392
Db      1196  GACAGCCTGACCCGTCGCCG---CACCTGTTCACCTGGCTGGACACGCTGAACCTTCAC 1252
Qy      393  -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
Db      1253  GAAAGACGCAACACCCCGCAACACTTCTTACAGCCCACTACAACATGTTCCACTAC 1312
Qy      402  ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
Db      1313  ACCCTGGACAACATCAGCCAAAGAGCGGTGTTGGCAACCAACACGACGACGACAAA 1372
Qy      421  LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440
Db      1373  CTG-----AAGAGCTGGGTCTGGCAACCAACATCTACATC----- 1408
Qy      441  LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
Db      1409  -----TTC 1411
Qy      461  IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
Db      1412  CTGCTGAACGTCACTAGCCTGGACACAAATACTTGAACGACTACAACAACATCAGCAAA 1471
Qy      477  -----LeuLysTyrSerAlaGlyGly 483
Db      1472  ATGGACTTCTTCATCAACACGGTACCGCCTGCTGGAAAAAGAACTGACCGCAGGCGAGC 1531
Qy      484  SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysLysAspCys 502
Db      1532  GGCAAAATCACCTACGACGTGAACAAACAAATCTTCGCTCGCATCTCGAAACGTCGC 1591
Qy      503  AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
Db      1592  GAAACCAAGGCAACCGACCTGTTCCGCACTACGACAACTACAGCACTCTCTGAGC 1651
Qy      523  HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThr 542
Db      1652  TTCATCAAAAGCTGACGCTCCGCAACCTACAAACCCCAAGTGATC----- 1699
Qy      543  GlyValLeuGlyTrpThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
Db      1700  ---ACCTTCGCGTGGACCCACGACGCTGACCCGAAAAACCACTCTACACCCACCTG 1756
Qy      563  IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582
```

```
Db      1757  ACCACCCAAATCCCGCGGTGAAGCGAAGCAGCTGGGCACCGGAGCAAAAGTGTCCTCAA 1816
Qy      583  GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle 602
Db      1817  GGTCCGGGTCAACACCGCGGTGACCTGATC-----GACTTCAAAGACCACTTCAAATC 1870
Qy      603  ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
Db      1871  ACCTGCCAACACACAGCAACTTCCAAACAAGCTACTTTCATCCGCATCCGTTAGCGAGCAAC 1930
Qy      623  GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db      1931  GGCAGCCCAACACCCCGCGGTGATCAACCTGAGCATCCCTGGCGTGGCAGAACTG--- 1987
Qy      643  ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
Db      1988  GGTATGGCACTGAACCCGACCTTCAGCGGTACCGACTACACCACTGAAATACAAAGAC 2047
Qy      663  PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
Db      2048  TTCCAATACCTGGAGTTTCAGCAACGAAGTGAATTCGGCGCGAACCACCAACATCAGCCTG 2107
Qy      683  IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleIleAspLysIleGluPhe 701
Db      2108  GTGTTCAACCGTAGCGAGCTGTACCAACACACCCGCTGCTGTGATCGACAAAATCGAGTTC 2167
Qy      702  IleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGln 721
Db      2168  CTGCCGATCACCCGTAGCATCCGCGAAGACCGTGAAAAACAAAACCTGGAACCGTGCA 2227
RESULT 11
AAQ14670
ID      AAQ14670 standard; DNA; 2061 BP.
XX      AAQ14670;
XX      27-AUG-2003 (revised)
DT      25-MAR-2003 (revised)
DT      04-FEB-1992 (first entry)
XX      Dipteran active toxin gene.
XX      Insecticide; B.t; crystal; delta endotoxin; cryIVC; ss.
XX      Bacillus thuringiensis serovar morrisoni.
XX      Key Location/Qualifiers
XX      CDS 1..2028
XX      /*tag= a
XX      EP457498-A.
XX      21-NOV-1991.
XX      09-MAY-1991; 91EP-00304180.
XX      15-MAY-1990; 90US-00524255.
XX      01-OCT-1990; 90US-00590903.
XX      (MYCO ) MYCOGEN CORP.
XX      Sick AJ;
XX      WPI; 1991-341902/47.
XX      P-PSDE; AAR14374.
XX      Bacillus thuringiensis genes encoding diptera-active toxins - and
XX      transformed microbes used to control insects in various environments.
XX      Claim 1; Page 15; 20pp; English.
XX      The sequence was obtd. from plasmid pMYC1636 which was isolated from a
XX      genomic library prepd. from DNA from B.t. PS71M3 [from B.t. PS71M3-69
CC
```

CC (NRRL B-18515)]. It is related to the cryIVC from B.t. var. israeliensis.
 CC The gene encodes a 77 kD protein. Microorganisms transformed with the DNA
 CC may be administered to dipteran insects or their environments, the
 CC expressed toxins acting as an insecticide. See also AAQ14669-Q14672.
 CC (Updated on 25-MAR-2003 to correct PA field.) (Updated on 27-AUG-2003 to
 CC correct OS field.)
 XX

SQ Sequence 2061 BP; 779 A; 331 C; 319 G; 632 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.86e-71	Length:	2061
Score:	899.50	Matches:	247
Percent Similarity:	48.32%	Conservative:	126
Best Local Similarity:	31.99%	Mismatches:	270
Query Match:	23.25%	Indels:	129
DB:	2	Gaps:	30

US-10-783-417-2 (1-735) x AAQ14670 (1-2061)

QY	1	MetAsnGlnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr	20
DB	1	ATGAATCCATATCAAAATAGATGATATGAATATTCATGCTCCATCCAAATGGTTTT	60
QY	21	PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro	40
DB	61	-----AGCAAGCTATAACTATTCTAGATATCCATAGCAATAAGCCAAATCAACCA	114
QY	41	LeuGlnAsnThrAsnTyrIleGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly	60
DB	115	CTCAAAACACGAATACAAAGATGGCTCAATGTGTGTCAGATAATCAACATATGGC	174
QY	61	AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIle	80
DB	175	AATAATCGCGGGAAATTTGTAGTTCTGAAACTATTCTGGAGTTAGTCAGGTATTATT	234
QY	81	ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly	100
DB	235	GTAGTAGGAAGTATGTAGGA-----GCTTTGCTGCCCT-----	270
QY	101	IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGly	120
DB	271	GTCTAGCTGCGGTATAATATCTTTGGGACTTTGTGGCGATCTTTTGG---CAAGGA	327
QY	121	GluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr	140
DB	328	TCTGACCCTGCAAATGTTTGGCAGGATTTCTTAAACATCGA-----GGAAGG	375
QY	141	ProLeuThrGlu-----SerIleGlyGlnLeuLysLeuGlnThrLeuGluGlyPhe	157
DB	376	CCATATCAAGAAATAGATAAAACATAAATTAATGTACTAATCTTCTATCGTAACACCTATA	435
QY	158	ArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrPArgLysLeuLysArgLeu	177
DB	436	AAAAATCAACTTGATAAATCAAGAAATTTTCGATAAATGGGAGCCGACGCGT-----	489
QY	178	GlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIle	197
DB	490	-----ACACAGCTAATCTTAAGACGATACATGATCTCTTTACT	528
QY	198	ArgPheGluAsnValHisAsn-----AspPheIleArgGluIleProGlyPheGln	214
DB	529	ACCTTAGAACCTATAATAGATAAAGATTTAGATATGTTAAATAAATAATGCTAGCTATCGA	588
QY	215	LeuGluTyrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeu	234
DB	589	ATACCAACA-----CTCCCTGCATATGCAAAATAGTACTTGTGCACCTTG	633
QY	235	AsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTyrPAsnAlaAspIleHisProSer	254
DB	634	AATTTATTAACATCGCTGCTACTTACATATATGG-----CTGCAAAATCAA	684
QY	255	GlnIleGluProAsnAlaGlyThrSerAspAspTyrTyrLys---LeuLeuLysGluAsn	273

DB	685	GGTATAAATCCAGTACTTTCAATTCATCTAATTACTATCAGGGCTATTATAAACGTAAA	744
QY	274	IleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArg	293
DB	745	ATAACAAGATATACTACTATTGTATACAAACGTACAAATGCAGACTACTATGATTAGA	804
QY	294	AspGluProAsnMetLysTyrSerIlePheAsnAspTyrArgArgTyrMetThrIleThr	313
DB	805	ACTAATACTAACGCAACATGGAATATGATATATCTTACCGTTTAGAATGACTCTTAAT	864
QY	314	ValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIle	333
DB	865	GTGTTAGATCTTATTGCTATTTCCTCAAAATATGATACCCAGAAATAT-----	912
QY	334	GlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrGlu	353
DB	913	-----CCAATAGGAGTTAAATCTGAACCTTACAGAGAAGTTTAT---ACGAAT	957
QY	354	IleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyr	373
DB	958	GTTAATTCAGATACATTT-----AGAACCATACAGAACTAGAAAAT	999
QY	374	AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr	393
DB	1000	GGATTAACCTAG---AATCCTACATTTATTCTTGATAAACCAAGGGCGTTTACACA	1056
QY	394	GluAsnThrAsn-----PheGlyAsnArgLeu	402
DB	1057	AGAAATTCGAGACATTTCTTGATCTTATGATATTTTCTTTTACAGGTAAACAGATG	1116
QY	403	ValGlyIleSerAsnAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyr	422
DB	1117	-----GCCTTTACACATCTAATGATGATCGCAACATAATCTCG	1155
QY	423	GlyGluArgThrGly-----SerProThrThrLysThrIleArgProPheGluSer	439
DB	1156	GGAGCGGTTCTAGGACATATTATTCTCAAGACATCCAAAGTATTTCTTTTATAGA	1215
QY	440	Tyr-----LysValSerIleValThrAspArgGlnSerProProValSerPro	455
DB	1216	AACAAACCTATTGTATAAGTCTGAAATTCGACAGATAGAGTACTCAGATATAATA---	1272
QY	456	IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn	475
DB	1273	-----TATGAAATGATATTTTTCGAAATAGCAGTGA	1305
QY	476	ThrLeuLysTyrSerAlaGlyGlySerLeu---SerAsnTyrGlnAsnThrThrPhePhe	494
DB	1306	GTATTTGATATTCATCCAAATTCACAAATAGAAAATAATTAATAAAGAACTGATTTAT	1365
QY	495	GlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPhe	514
DB	1366	ATGATTCGAAACAAACATCGGAAAT-----	1392
QY	515	AsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGly	534
DB	1393	AAAGAAATATGCTACTCTATCGTATATAAATACTGATAATATATATATTTTTCAGTAGT	1452
QY	535	LeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArg	554
DB	1453	AGAGAAAGAAAGAGATTT-----GCATTTAGTTGGACACATACATAGTGTGTTTC	1503
QY	555	TyrAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnLeu	574
DB	1504	CAAAATACATAGATTTAGATAACATCCCAATTCACGCTCTAAAGCTTTGAAGGTA	1563
QY	575	AspThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeu	594
DB	1564	AGTTCTGATCGAAATTTGTGAAAGGCTCTGTGCACAGGTGGAGACTTGTGTAATCTTT	1623
QY	595	GlnSerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPhe	614
DB	1624	AAAGATAGTAGGATTTTAGAGTTAGATTT---TTAAAAAATGTTTTCGACAATACAA	1680

QY 615 IleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThr 634
Db 1681 GTACGATTGTTGTTACTAATGCT- - - - - CCAAGACAAACAGTATTC 1725
QY 635 IleProGlyValIleGlyIleProGlnArgLeuAsnAsnThrPheSerGlyThrAsn 654
Db 1726 TTAAACCGAATAGATATAAGTGGAG- - - CTCCCTAGTACCCTCCGCCAATAC 1782
QY 655 TyrAsn- - - - - AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThrVal 672
Db 1783 CCAAAATGCTACAGATTTAAACATATGAGATTTGGATATGATCAATTTTCCAAGAACAGTT 1842
QY 673 - - - - - ThrLeuProLeuAsnArgAsnIleProPheIle 683
Db 1843 CCAAAATAAACATTTGAAGGAGAACACTTTTAAATGACC- - - - - TTA 1887
QY 684 PheAsnArgAlaAspValSerAsnSerIleLeuIleAspLysIleGluPheIlePro 703
Db 1888 TATGGTACACCAAAATCATTTATATATATA- - - TATATGACAAAATCGAAATTTATTTCCA 1944
QY 704 IleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLys 723
Db 1945 ATCACTCAATCTGTATTAGATTATACAGACAGCAAAATATAGAAAAACACAGAAATA 2004
QY 724 IleAsnThrPheThrAsnHisThrLysThrLeu 735
Db 2005 GTGAATGATTATTATTGTTAATTAACAAAGTTCTT 2040

RESULT 12
AAQ81180
ID AAQ81180 standard; DNA; 2061 BP.
XX
AC AAQ81180;
XX
DT 25-MAR-2003 (revised)
DT 12-AUG-1995 (first entry)
XX
DE B.t. toxin PS71M3 gene.
XX
KW Delta-endotoxin; crystal protein; biological control agent; Calliphorid;
KW screw-worm; sheep blowfly; Lucilia; Phormia; Calliphora; insecticide;
KW pesticide; B.t; ss.
XX
OS Bacillus thuringiensis.
XX
PN WO9502694-A2.
XX
PD 26-JAN-1995.
XX
PF 13-JUL-1994; 94WO-US007902.
XX
PR 15-JUL-1993; 93US-00093199.
XX
PA (MYCO) MYCOGEN CORP.
XX
PI Hickie LA, Payne J;
XX
XX
DR WPI; 1995-067338/09.
DR P-PSDB; AAR63079.
XX
PT Method for controlling Calliphoridae pests - specifically utilises
PT Bacillus thuringiensis isolates or toxins.
XX
PS Disclosure; Page 42-43; 50pp; English.
XX
CC A library was constructed from Bacillus thuringiensis PS71M3 total
CC cellular DNA in lambda Gem-11. Plasmid pMYC1636, selected in Escherichia
CC coli, contained a 15 kb insert expressing a beta-endotoxin gene. This was
CC sequenced (AAQ81180). A cured, acrystalliferous B.t. host carrying
CC pMYC1636 produced a 77 kDa crystal protein (AAR63079). (Updated on 25-MAR
CC -2003 to correct PN field.)
XX
XX

SQ Sequence 2061 BP; 779 A; 331 C; 319 G; 632 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3,86e-71 Length: 2061
Score: 899.50 Matches: 247
Percent Similarity: 48.32% Conservative: 126
Best Local Similarity: 31.99% Mismatches: 270
Query Match: 23.25% Indels: 129
DB: 2 Gaps: 30
US-10-783-417-2 (1-735) x AAQ81180 (1-2061)
QY 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 1 ATGAATCCATATCAAAATGAATGAATGAATATTCATATGCTCCATCAATCGTTT 60
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 61 - - - - - AGCAAGTCTAATACTATTCTAGATATCCATTAGCAATAGCCAAATCAACA 114
QY 41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
Db 115 CTGAAACACGAAATACAAAGATTGCTCAATGCTCAAGATAATCAACAATATGCG 174
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIle 80
Db 175 AATAATCGGGGAATTTTGTAGTTCTGAAACTATTGTGAGATTAGTCAGGTATTATT 234
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGly 100
Db 235 GTAGTAGAACTATGTTAGGA- - - - - GCTTTGCTGCCCT- - - - - 270
QY 101 IleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120
Db 271 GTCTTAGCTGAGGTATAATATCTTTGGGACTTTGTTGCCGACTTTTGG- - - CAAGA 327
QY 121 GluGlnAspLysThrValThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db 328 TCTGACCTGCAAAATGTTGGCAGGATTTGTTAAACATCGGA- - - - - GGAAGG 375
QY 141 ProLeuThrGlu- - - - - SerIleGlnLeuLysLeuGlnThrLeuGluGlyPhe 157
Db 376 CCTATACAGAAATAGATAAAACATAATTAATGCTACTTAATCTTCTGTAACACCTATA 435
QY 158 ArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeu 177
Db 436 AAAAAATCACTTGATAAATATCAAGAAATTTTCGATAAATGGAGCCAGCAGT- - - - - 489
QY 178 GlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIle 197
Db 490 - - - - - ACACAGCTAAATGCTAAAGCATACATGATCTCTTTACT 528
QY 198 ArgPheGluAsnValHisAsn- - - - - AspPheIleArgGluIleProGlyPheGln 214
Db 529 ACCTTAGAACCTATAATAGATAAAGATTAGATATGTTAAATAATAATGCTAGCTATCGA 588
QY 215 LeuGluThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeu 234
Db 589 ATACCAACA- - - - - CTCCCTGCATATGCACAAATAGCTACTTGGCAGCTTG 633
QY 235 AsnLeuLeuGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSer 254
Db 634 AATTTATTAACACATGCTGTACTTACCTATTACAATATATG- - - - - CTGCAAAATCAA 684
QY 255 GlnIleGluProAsnAlaGlyThrSerAspAspTyrTyrLys- - - - - LeuLeuLysGluAsn 273
Db 685 GGTATAAATCCAGTACTTTCAATTCATTAATTAATCTATCAGGGCTATTTAAAAACGTAAA 744
QY 274 IleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArg 293
Db 745 ATACAGAAATATCTGACTATTGTATACAAACGTAACATGATGAGGACTATGATGATTAGA 804
QY 294 AspGluProAsnMetLysTrpSerIlePheAsnAspTyrArgTyrMetThrIleThr 313


```
Db      805 ACTAATACTAAACGCAACATCGAATATCTATAATACTTACCGTTTAGAAATGACTCTAACT 864
Qy      314 valLeuaspThrIleSerGlnPheSerLeuTyrAspIleYsArgTyrArgAspSerIle 333
Db      865 GTGTTAGATCTTATGCTATTTTCCAAATATATGACCCAGAAATATAT- 912
Qy      334 GlyGlyIleGluValIysGlyIleLeuAsnGluLeuThrArgGluIleTyrThrGlu 353
Db      913 -----CCATAGAGAGTTAAATCGAACTTACGAGAGAGTTTAT---ACGAAT 957
Qy      354 IleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyr 373
Db      958 GTTAATTCAGATACATTT-----AGAACCAATAACAGAACTAGAAAAT 999
Qy      374 AsnLeuThrArgAlaSerPheLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
Db      1000 GGATTAACCTAGA---AATCCTACATTTATTTCTGGATAAACCAAGCGCGTTTTTACACA 1056
Qy      394 GluAsnThrAsn-----PheGlyAsnArgLeu 402
Db      1057 AGAAATTCGAGACATCTTGATCTTATGATATTTTCTTTTACAGGTAACCAATG 1116
Qy      403 ValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyr 422
Db      1117 -----GCCTTTACATACTAATGATGATGATCCACATATATCTGG 1155
Qy      423 GlyGluArgThrGly-----SerProThrThrIysThrIleArgProPheGluSer 439
Db      1156 GGAAGCGGTTTCATCGACATATTTTCTCAAGACACATCCAAAGTATTTCTTTTATAGA 1215
Qy      440 Tyr-----LysValSerIleValThrAspArgGlnSerProProValSerPro 455
Db      1216 AACAAACCTATTGATAAGGTCGAAATTTGTGACATAGAGAGTACTCAGATATAATA- 1272
Qy      456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn 475
Db      1273 -----TATGAATGATATTTTTTTCGAATAGCAGTGAA 1305
Qy      476 ThrLeuLysTyrSerAlaGlyGlySerLeu---SerAsnTyrGlnAsnThrThrPhePhe 494
Db      1306 GTATTTTCGATATTCATCCATTCACACATAGAAAATATTTATAAAGAACTGATCTTAT 1365
Qy      495 GlnPheProArgLysIysAspCysAsnLeuValIleAspProGlyCysSerProAsnPhe 514
Db      1366 ATGATTTCCAAAACAAACATCGAAAAAT----- 1392
Qy      515 AsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGly 534
Db      1393 AAAGAATATGGTCATATCTCTATCTATATAAATACTGATAATATATTTTTCAGTAGTT 1452
Qy      535 LeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArg 554
Db      1453 AGAAGAAGAGAGAGAGTT-----GCATTTAGTTGGACACATACTAGTTGATTC 1503
Qy      555 TyrAsnAlaIleSerAspIysIleIleThrMetIleProAlaIleIysGlyAsnAsnLeu 574
Db      1504 CAAAATACAAATAGATTAGATAACATCACCCAAATCCACGCTCTAAAGGCTTGAAGGTA 1563
Qy      575 AspThrAsnSerIysValIleGluGlyProGlyHisThrGlyAsnLeuValTyrLeu 594
Db      1564 AGTCTCTGATTGCGAAAAATTTGAAAGGTCCTGGTCACACAGGTGAGACTTGGTAATTCCT 1623
Qy      595 GlnSerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPhe 614
Db      1624 AAAGATAGTAGTATGATTTTAGAGTTAGATTT---TTAAAAATGTTTCTCGACATATCA 1680
Qy      615 IleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThr 634
Db      1681 GTACGTTATCGTTATGCTACTATGCT-----CCAAAGACAAACAGTATTC 1725
Qy      635 IleProGlyValIleGlyIleProGlnArgLeuAsnAsnThrPheSerGlyThrAsn 654
```

```
Db      1726 TTAACCGGAATAGATACTATTAAGTGTGGAG---CTCCCTAGTAGCACCTTCCCGCCAAAC 1782
Qy      655 TyrAsn-----AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThrVal 672
Db      1783 CCAAAATGCTACAGATTTTAAACATATGCAATTTTGGATATGTAACATTTTCCAAGACAGTT 1842
Qy      673 -----ThrLeuProLeuAsnArgAsnIleProPheIle 683
Db      1843 CCAAAATAAACATTTGAAGGAGAGACACTTTATTAAATGACC-----TTA 1887
Qy      684 PheAsnArgAlaAspValSerAsnSerIleLeuIleAspLysIleGluPheIlePro 703
Db      1888 TATGGTACACCAATCATTCATATAATA---TATATTGACAAAATCGAATTTATTTCCA 1944
Qy      704 IleThrSerSerMetHisGlnAsnArgGlnLysGlnLysLeuGluThrIleGlnThrLys 723
Db      1945 ATCACTCAATCTGTATTAGATTATACAGAGAAGCAAAATATAGAAAAAACACAGAAAAATA 2004
Qy      724 IleAsnThrPhePheThrAsnHisThrLysThrLeu 735
Db      2005 GTGAATGATTTATTGTTTAAATTAACAAAGTTCTT 2040

RESULT 13
AAN50525
ID AAN50525 standard; DNA; 3756 BP.
XX AAN50525;
AC AAN50525;
XX XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 23-OCT-1991 (first entry)
XX XX
DE Bacillus thuringiensis var. israelensis endotoxin insert in plasmid
DE pSY367.
XX XX
KW Endotoxin; insecticide; ss.
XX XX
OS Bacillus thuringiensis serovar israelensis.
XX XX
PN EPI53166-A.
XX XX
PD 28-AUG-1985.
XX XX
PF 15-FEB-1985; 85EP-00301017.
XX XX
PR 22-FEB-1984; 84US-00582506.
PR 22-JAN-1985; 85US-00693556.
XX XX
PA (SYTR ) SYNTRO CORP.
XX XX
PI Walfield AM, Pollock TJ;
XX XX
DR WPI; 1985-211724/35.
XX XX
PT Polypeptide active against Diptera insects - prep'd. from DNA sequence
PT coding for BTI endotoxin using bacterial host.
XX XX
PS Disclosure; Page 15a-e; 27pp; English.
XX XX
CC The B. thuringiensis var. israelensis endotoxin insert in pSY367 is
CC expressed in a bacterial host. The protein produced has insecticidal
CC activity against dipteran insects. (Updated on 25-MAR-2003 to correct PA
CC field.) (Updated on 24-OCT-2003 to standardise OS field)
XX XX
SQ Sequence 3756 BP; 1379 A; 566 C; 659 G; 1150 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.: 4.2e-69 Length: 3756
Score: 881.00 Matches: 248
Percent Similarity: 47.42% Conservative: 120
Best Local Similarity: 31.96% Mismatches: 272
Query Match: 22.77% Indels: 136
DB: 1 Gaps: 31
```


QY 684 -----PheAsnAlaAspValSerAsnSerIleLeuIleAspIle 699
 Db 2819 AATGACCTTATATGTGTACACCAATCATATATATA---TATATGACAAAT 2875
 QY 700 GluPheIleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuThr 719
 Db 2876 GAATTTATTCATCACTCAATCTGTATTAGATTATACAGAGAAGCAAAATATAGAAAA 2935
 QY 720 IleGlnThrLysIleAsnThrPhePheThrAsnHisThrLysThrLeu 735
 Db 2936 ACACAGAAATAGTGAATTTATTTGTTAATTAACAAAGTTCTT 2983

RESULT 14

AAD43974

ID AAD43974 standard; DNA; 4896 BP.

XX AAD43974;

AC AAD43974;

DT 13-DEC-2002 (first entry)

XX Bacillus thuringiensis ssp. finitimus cry28Aa1 gene.

XX Delta-endotoxin; cry26Aa1; cry28Aa1; insect-resistant plant; toxin;

XX transgenic host cell; insecticide; gene; ds.

XX Bacillus thuringiensis.

FH Key Location/Qualifiers

FT 1129..4458

FT /*tag= a

FT /product= "cry28Aa1 gene"

XX US2002038005-A1.

XX 28-MAR-2002.

XX 08-JAN-2001; 2001US-00756526.

XX 07-JAN-2000; 2000US-0175158P.

XX (WOJC/) WOJCIECHOWSKA J A.

XX (LEWI/) LEWITIN E I.

XX (ZALU/) ZALUNIN I A.

XX (REVI/) REVINA L P.

XX (CHES/) CHESTUKHINA G G.

XX Wojciechowska JA, Lewitin EI, Zalunin IA, Revina LP;

XX Chestukhina GG;

XX WPI; 2002-403936/43.

XX P-PSDB; AAB26353.

XX Novel isolated delta-endotoxin nucleic acid molecules, cry26Aa1 and

XX cry28Aa1 isolated from *Bacillus thuringiensis finitimus*, that encodes

XX toxin active against insects, useful for controlling insects.

XX Claim 1; Page 33-37; 42pp; English.

XX The invention relates to isolated delta-endotoxin nucleic acid molecules,

XX cry26Aa1 and cry28Aa1 isolated from *Bacillus thuringiensis finitimus*,

XX that encode a toxin that is active against insects. The invention is

XX useful for producing an insect-resistant plant, by introducing the

XX nucleic acid molecule into the plant, where the nucleic acid is

XX expressible in the plant in an effective amount to control an insect. The

XX invention is useful for producing a toxin that is active against insects

XX by obtaining the transgenic host cell and expressing the nucleic acid

XX molecule in the host cell, which results in the toxin that is active

XX against insects. The toxin is useful for controlling an insect by

XX delivering to the insect an effective amount of toxin. The invention is

XX useful for controlling insects. The toxin is useful for inhibiting the

XX ability of insect pest to survive, grow or reproduce, for limiting insect

XX related damage or loss in crop plants, and to prophylactically treat

CC insect susceptible areas or plants to confer protection or resistance
 CC against harmful insects. The present sequence is *Bacillus thuringiensis*
 CC ssp. *finitimus cry28Aa1* gene

XX
 SQ Sequence 4896 BP; 1727 A; 757 C; 887 G; 1525 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4,8e-66 Length: 4896
 Score: 849.00 Matches: 229
 Percent Similarity: 48.76% Conservative: 106
 Best Local Similarity: 33.33% Mismatches: 248
 Query Match: 21.94% Indels: 104
 DB: 6 Gaps: 25

US-10-783-417-2 (1-735) x AAD43974 (1-4896)

QY 68 SerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAla 87
 Db 1210 AGCAGTGATACAGTCGCTGTAGTAGCGCAGGGATTGATTGGGTACTATCTGACA 1269
 QY 88 GlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIle 107
 Db 1270 -----GCCTTTGCATCATTTGTTAATCCA-----GGTGGGTACTATATA 1308
 QY 108 SerPheGlyThrLeuIleThrValPheThrProAlaGlyGluGlnAsp---LysThrVal 126
 Db 1309 TCATTGGAACTTGGCTCCGCTTTGGGCTGATCAGAGGAGATCCAAAAAAT 1368
 QY 127 TrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeuThrGluSerIle 146
 Db 1369 TGGTCACATTTATGAAACACGAGACACCTTTTAATCAACAATTTCTACAGCTGTA 1428
 QY 147 LysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGlnSerTyrAsnThr 166
 Db 1429 AAGAAATAGCATTAGCTCATCTAAATGCTTTAAAGATGATTAAACGCTATGAAAGA 1488
 QY 167 AlaLeuAspAspThrArgLysLeuLysArgLeuGlnAlaProGlyLeuProSerSer 186
 Db 1489 GCATTTAATGATTGGAAGAGA-----AATCCAAGTGCA 1521
 QY 187 AlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPhe 206
 Db 1522 -----AATCTGCCAGATTGGTATCACAGAGATTGAAACGCTCATTTCAATTT 1572
 QY 207 IleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeuLeuProIleTyr 226
 Db 1573 GTAAGCAATATGCCACAACCTCCCACTCCCAAGTATGACACATATTATTAAGTCTAT 1632
 QY 227 AlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGlu 246
 Db 1633 ACAGAAGCTGCNAATTTACATTTGAAATTTATATACATCAAGGTGTACAAATTCGGGATCAA 1692
 QY 247 TrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTyr 266
 Db 1693 TGGAAATCAGATCAACACCATTCACCAATGTGTAAGTCATCAGGTACT-----TAT 1743
 QY 267 TyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyrArg 286
 Db 1744 TATCAGCAGCTATTTGGTATATATTGAAAGATATATTATTTATGCACCAAGACATACCAT 1803
 QY 287 ThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTyr 306
 Db 1804 AAGAGTTGAATCACTTAAAGAAATCAGAAAAATCATATGGGATGCTTTATAACACATAT 1863
 QY 307 ArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIle 326
 Db 1864 CGTCGAAATAGACCTTAATTTGATTGGAATCTTGTCCCAACTTTTCTTTTATGATATA 1923
 QY 327 LysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThr 346
 Db 1924 CGTCGTTTT-----CCAGAGGAGGTAGAACTAGAAATTAACA 1959
 QY 347 ArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeuArgValGlnPro 366

Query Match:	21..94%	Indels:	104
DB:	10	Gaps:	25
US-10-783-417-2 (1-735) x ADF31301 (1-4896)			
Qy 68 SerAlaAspThrIleAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAla 87			
Db 1210 AGCAGTGCATACAGTCGCTAGTAGTAAGCGCAGGAGTGTAGTGGGTACTACTGACA 1269			
Qy 88 GlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIle 107			
Db 1270 -----GCCTTGTGCATCATTTGTTAAATCCA-----GGTGTGGTACTTATA 1308			
Qy 108 SerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGlnAsp---LysThrVal 126			
Db 1309 TCATTTGGAAACCTTGGCTCCGTTCTTGGCTGATCCAGAGGAAGATCCAAAAAAT 1368			
Qy 127 TrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeuThrGluSerIle 146			
Db 1369 TGGTCACAATTTATGAACACGAGAGACCTTTTAAATCAACAATTTCTACAGCTGTA 1428			
Qy 147 LysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGlnSerTyrAsnThr 166			
Db 1429 AAAGAAATAGCATTAGCTCATCTAAATGGTTTAAAGATGTATTAAAGCTACTATGAAGA 1488			
Qy 167 AlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeuProProSerSer 186			
Db 1489 GCATTTAATGATTGGAAGAGA-----AATCCAAGTGCA 1521			
Qy 187 AlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPhe 206			
Db 1522 -----AATACTGCCAGATTGGTATCACAGAGATTGAAACGCTCATTTCAATTTT 1572			
Qy 207 IleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeuLeuProIleTyr 226			
Db 1573 GTAAGCAATATGCCAACCTCCAACTTCCACCGTATGCACATTTATTATTAAGTTGTAT 1632			
Qy 227 AlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlnGlyAlaGluLeuAlaAspGlu 246			
Db 1633 ACAGAAGCTGCAAAATTTACATTTGAATTTATTATACATCAAGGTGPACAATTCGCGGATCAA 1692			
Qy 247 TrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTyr 266			
Db 1693 TGAATGTCAGATCAACCACTTACCAGTGTGAAGTCATCAGGTACT-----TAT 1743			
Qy 267 TyrLysLeuLysGluAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyrArg 286			
Db 1744 TATGACGAGCTATTGGTATATATTGAAAAGTATATTATTATTTGCACCAAGACATACCAT 1803			
Qy 287 ThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTyr 306			
Db 1804 AAAGAGTTGAATCACCTTAAAGAAATCAGAAAAATCATCATGGGATGCTTATAACACATAT 1863			
Qy 307 ArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIle 326			
Db 1864 CGTCGAGAAATGACCTTAATTTGTTATGCGATCTTGTGCGCACTTTCTCTTTTATGATATA 1923			
Qy 327 LysArgTyrArgAspSerIleGlyIleGluValLysGlyIleLysAsnGluLeuThr 346			
Db 1924 CGTCGTTT-----CCAAGAGGAGTAGAACTAGAAATTAAACA 1959			
Qy 347 ArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeuArgValGlnPro 366			
Db 1960 AGAGAGTTTATACAGTTTATAGATCATTTAAACAGACCA----- 2001			
Qy 367 AsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeu 386			
Db 2002 -----GGGCTATTACTTGGCTG 2019			
Qy 387 GluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIleSer 406			
Db 2020 TCAGATATTGAGTTATACACGGAGAGTGTGGCAGAGCGGATTTATTATTCAGGTATT--- 2076			
Qy 407 AsnArgAspAlaProThrTyrSerAsnThr-----IleThrGluThrLeuTyrGly 423			
Db 2077 ---CGAGAGTCTAAATATTACTGTAATCAATTTTTCACGATGAAATAAATATTATGTT 2133			
Qy 424 GluArgThrGly---SerProThrThrLysThrIleArgProPheGluSerTyrLysVal 442			
Db 2134 AATACAAATAGATTAAAGTAAGCAGCTCATTTATTTACACAGCGAA---TTTATGACT 2190			
Qy 443 SerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIleIle 462			
Db 2191 CACTTAAGCATAAACCGCTCTTTTCAACAATAGCTGGTATATAAATAGTTATACAGTTTA 2250			
Qy 463 AsnGlnIleGluLeuThrLeuAsnGlySerSerAsnAsnThrLeuLysTyrSerAlaGly 482			
Db 2251 ATTCAAAAAATCGTATTTCACAACTTTTAAACAGATAAT-----GAATATCAAAAAAT 2304			
Qy 483 GlySerLeuSerAsnTyr-----GlnAsnThrThrPheGlnPheProArgLys 499			
Db 2305 TTTAATGTGAATAATCAAAATGAACCTCAAGAAACTACA----- 2343			
Qy 500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHis 519			
Db 2344 -----AACTATCTCTAATGATTTATGTTGTTCA---AACAGCCAAAAATTCACACAT 2391			
Qy 520 IleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIle 539			
Db 2392 AATTATCTCTCAATTTCCATTAATCATCCAC-----AAGTTAGATTTT 2433			
Qy 540 LeuAsp-----ThrGlyValLeuGlyTrpThrHisSerSerValAspArgTyr 555			
Db 2434 GCTGAGTATTTTCACTCTATATTGTCATAGTTGGACACACAAATAGTGTAACTCCCAA 2493			
Qy 556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp 575			
Db 2494 AATTAAATATCAGAAAGTGTGAGTACACAAATCCCATTTGGTAAAGCTTTACGAAGTT--- 2550			
Qy 576 ThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln 595			
Db 2551 ACTAACAAATTCAGTTATTAAGAGAGCAGGTTTACAGGTGGAGATTAAATAAGAACTCGT 2610			
Qy 596 SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIle 615			
Db 2611 GAT-----AAATGTTCTATTAAATGTAAGCT---AGTTCTTTTAAAAAATACGCTATA 2661			
Qy 616 ArgLeuArgTyrAlaThrAsnGlyValaGlyAsnThrLeuProAsnIleSerLeuThrIle 635			
Db 2662 AGTCTATTTTATGCTGCAAAATACGCAATAGCTGTATCAATAGACGTAGGTGATCCGGA 2721			
Qy 636 ProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyr 655			
Db 2722 GCAGGAGTTCTA-----TTGCAACCTACCTTTTCTAGAAAAAGGGAAC 2763			
Qy 656 AsnAsn-----LeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSer 670			
Db 2764 AATAATTTTACAATTCAGACCTTAACATTAAGGATTTTCAATATCATACACTTTTAGTT 2823			
Qy 671 ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAsp----- 688			
Db 2824 GATATTGAAATTCACGAAAGTGAAGAAATTCATATCCATTTGAACGGAGAGATGATAT 2883			
Qy 689 ValSerAsnSerIleLeuIleAspLysIleGluPheIleProIleThrSerSerMet 708			
Db 2884 GAGGAGGAGTGATCTCTTTTAAATGATAAATAGAGTTTCAAACTTATAGATGAAATAT 2943			
Qy 709 HisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePhe 728			
Db 2944 -----ACTAATGAATGAATTTAGAGAAGGCAAGAAAGACGTGAATGTATTATTT 2994			
Qy 729 ThrAsnHisThrLysThrLeu 735			
Db 2995 ATAAACGCAACAAACGCTTTG 3015			

Search completed: December 11, 2005, 23:38:24
Job time : 1173 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: December 11, 2005, 23:17:19 ; Search time 7267 Seconds
(without alignments)
4732.146 Million cell updates/sec

Title: US-10-783-417-2
Perfect score: 3869
Sequence: 1 MNQNDNNEVEIIDSHTSPV.....KLETIQTKINTFFNHTKTL 735

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 2339354128 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10783417/runat_01122005_141957_24358/app_query.fasta_1.903
-DB=EST -OFMT=fastap -SUFFIX=ret -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10783417 @CGN 1.1 5315 @runat 01122005_141957_24358 -NCPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCKS=100 -LONGLOG
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:**

1: gb_est1:**
2: gb_est2:**
3: gb_est3:**
4: gb_est4:**
5: gb_est5:**
6: gb_est6:**
7: gb_est7:**
8: gb_est8:**
9: gb_est9:**
10: gb_est10:**
11: gb_est11:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	147.5	3.8	6567	11	DQ031803 Homo sapi
2	124	3.2	13107	11	DQ035722 Homo sapi
3	121.5	3.1	3477	10	CL977084 OsiFCC029
4	121	3.1	1754	10	AG392993 Mus muscu
5	121	3.1	2052	4	CNS0A8VH
6	119.5	3.1	6777	10	CL967047 OsiFCC014
7	117.5	3.0	1986	10	AY416726 Mus muscu

8	117	3.0	2157	4	AK089255	Mus muscu
9	115	3.0	1178	1	AJ538005	AJ538005
10	115	3.0	2322	10	CL960568	OsiFCC004
11	115	3.0	2460	4	CNS09Y2	XB832541 Arabidops
12	114.5	3.0	820	7	CF951415	CF951415 UI-M-HL0-
13	112.5	2.9	1123	10	AG378595	AG378595 Mus muscu
14	112.5	2.9	2841	10	CL974457	CL974457 OsiFCC025
15	111	2.9	1143	4	CNS0A1JE	XB829715 Arabidops
16	111	2.9	1616	10	AG396120	AG396120 Mus muscu
17	110.5	2.9	1105	3	BM463012	BM463012 AGENCOURT
18	109.5	2.8	752	7	CO096813	CO096813 GR_Ea20E
19	109.5	2.8	861	7	CO070923	CO070923 GR_Ea28D
20	109	2.8	1878	2	BF528790	BF528790 602041401
21	108.5	2.8	949	7	CV674212	CV674212 RET781_30
22	108.5	2.8	956	4	AY812233	AY812233 Schistoso
23	108.5	2.8	2282	4	AK044734	AK044734 Mus muscu
24	108.5	2.8	5001	4	CR860285	CR860285 Pongo pyg
25	107.5	2.8	2223	10	CL981417	CL981417 OsiFCC045
26	107.5	2.8	2439	10	CL982610	CL982610 OsiFCC048
27	107.5	2.8	2955	10	CL947058	CL947058 OsiFSSB002
28	107	2.8	859	9	AZ548115	AZ548115 ENTDL91TR
29	106.5	2.8	1541	4	BC015498	BC015498 Homo sapi
30	106.5	2.8	1598	4	BC030976	BC030976 Homo sapi
31	106.5	2.8	2355	10	AY413036	AY413036 Mus muscu
32	106.5	2.8	2948	4	CNS0A4RI	XB822972 Arabidops
33	106.5	2.8	4751	4	BR859511	BR859511 Pongo pyg
34	106	2.7	881	9	HH162188	HH162188 ENTRY54TR
35	106	2.7	910	9	AZ546690	AZ546690 ENTEN48TF
36	106	2.7	2520	10	AG280133	AG280133 Mus muscu
37	106	2.7	3054	10	CL974410	CL974410 OsiFCC025
38	105.5	2.7	1092	10	CL065922	CL065922 CH216-105
39	105	2.7	622	7	CV176047	CV176047 SPHDROMA
40	105	2.7	697	7	CV533100	CV533100 LVS_035_B
41	105	2.7	747	3	BJ311235	BJ311235 BJ311235
42	105	2.7	16387	11	DQ038517	DQ038517 Pan trogl
43	105	2.7	16425	11	DQ038516	DQ038516 Homo sapi
44	104.5	2.7	458	1	AU006750	AU006750 AU006750
45	104.5	2.7	872	6	CD792427	CD792427 EST663788

ALIGNMENTS

RESULT 1	DQ031803	Homo sapiens	6567 bp	DNA	linear	GSS 02-JUN-2005
LOCUS	DQ031803	Homo sapiens	HC2093 gene, VIRTUAL TRANSCRIPT, partial sequence,			
DEFINITION	DQ031803	genomic survey sequence.				
ACCESSION	DQ031803	DQ031803.1	GI:56883012			
VERSION	DQ031803	GSS.				
KEYWORDS	DQ031803	Homo sapiens (human)				
SOURCE	DQ031803	Homo sapiens				
ORGANISM	DQ031803	Homo sapiens				
REFERENCE	DQ031803	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	DQ031803	1 (bases 1 to 6567)				
TITLE	DQ031803	Hubisz, M.J., Fladel-Alon, A., Tanenbaum, D.M., Civallo, D., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.				
JOURNAL	DQ031803	A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees				
PUBLISHED	DQ031803	(er) PLOS Biol. 3 (6), E170 (2005)				
REFERENCE	DQ031803	15969325				
AUTHORS	DQ031803	2 (bases 1 to 6567)				
TITLE	DQ031803	Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B., Hubisz, M.J., Fladel-Alon, A., Tanenbaum, D.M., Civallo, D., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.				
JOURNAL	DQ031803	Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA				
COMMENT	DQ031803	This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.				

```

FEATURES
  source      Location/Qualifiers
 1..6567      /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
              /chromosome="12"
  gene        /1..>6567
              /locus_tag="HC2093"

ORIGIN
Alignment Scores:
Pred. No.:    0.000318      Length:    6567
Score:        147.50        Matches:  160
Percent Similarity: 33.63%      Conservative: 101
Best Local Similarity: 20.62%    Mismatches: 238
Query Match:   3.81%          Indels:    277
DB:           11              Gaps:     40

US-10-783-417-2 (1-735) x DQ031803 (1-6567)
QY 10 TyrGluIleIleAspSerHisThrSerPro---TyrPheProAsnArgAsnSerAsnAsp 28
Db 1951 TATGAAGTCTGTATATAAATATAGATACCTTATATATAGAAACACATCAACACAGAC 2010
QY 29 -----SerArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsnThr 44
Db 2011 ATAATATTAAAGGAACCTTAAGACCTCACAC-----CTCTATAACATT 2052
QY 45 AsnTyrLysGluTrpLeuAsnMetCysGlnGlyAsnThrGlnTyrGlyAspAsnPheGlu 64
Db 2053 TCTGTAAAGTCTTACACACAGATTGGTTCATGGCAATCAGGTA-----TCTTCTTTACTC 2106
QY 65 ThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSerGlyThr 84
Db 2107 TCTGTAAAGACTTCGGAGACTGTGCCGTAGTAGCCACGAAATATACACTACAAAAAT 2166
QY 85 LeuLeuAlaGlyIleGlyLeuThrSerIle-----SerGlyProIleGlyIleIle 102
Db 2167 ATTTCCTCTGGAGAGATTGAGCTATCATCTCTCCCAAGTAGTCCCAATGGAATCATA 2226
QY 103 GlyAlaIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGln 122
Db 2227 AAAAAATATACAAATTTATCTCAAGAGA-----AGTAATGGAAATGAG 2268
QY 123 AspLysThrValTrpThrGlnPheIleLysMetGlyLullePheValAspThrProLeu 142
Db 2269 GAAAGAACTATAAATACA-----ACCTCTTTA 2295
QY 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeu-----154
Db 2296 ACCCAAAACATTAAAGTACTGAGAAATATACCCNATATATCATTTAGGTGTCTGTAGT 2355
QY 155 -----GluGlyPheArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAsp 170
Db 2356 ACACCTGAAAGGTGAAGGAGTTCGGAGTGTCCCAATAGTATCTGACGGGGAAGAT--- 2412
QY 171 TrpArgLysLeuLysArgLeuGlnAlaProGlyLeuProProSer-----SerAlaLeu 188
Db 2413 -----GCTCCGTGATTCCTCCCTCAAGACTTCTCTGTAAAA 2448
QY 189 GlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPheIleArg 208
Db 2449 CAGTTCTGGTGTACGGTGAAGTGTGTCTATGGCAA-----2484
QY 209 GluIleProGlyPheGln-----LeuGluThrTyrLysThrLeuLeuLeuProIle 225
Db 2485 -----CCACCCCTGGAGCCAAATGGAATTATCTTTATTACAGTTTAT-----GTC 2532
QY 226 TyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAsp 245
Db 2533 TGGAAATAGATCATCATTAATAAACTATTAAATGTCACTGAAACATCATTTGGAGTTATCAGAT 2592
QY 246 -----GluTrpAsnAlaAspIleHisProSer-----254

```

```

Db 2593 TTGATTATATAATGTGTAATACAGTCTTATGTAACAGCTAGCAGAGATTGTGGTGATGGG 2652
QY 255 -----GlnIleGluProAsnAlaGlyThr 262
Db 2653 AAAACAAGAACAATATATCAATTAGCTTTCAAACACAGAGGGAGCACCAGCGATCTCTCCC 2712
QY 263 SerAspAspTyrTyrLysLeuLeuLysGluAsn-----273
Db 2713 AAAGATGTTTATTATGCAAACTTCAGTTCTTCATCAATAAATCTTTTTCGGACACTCTCT 2772
QY 274 -----IleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThr 287
Db 2773 TCAAAACCTAATGGGATTATACAATATTAATCTCTGTTTATTACAGAAATACTTTCAGGTACT 2832
QY 288 GlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTyrArg 307
Db 2833 TTTATGCAGNNNNNNNNNTCCAT-----GAAGTAAACCAATGACTTTGAC 2877
QY 308 ArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIle--- 326
Db 2878 AATATGACTGTATCCACAATATATAGATAAACTGACAATATTTCAGCTACTATATACATTTTGG 2937
QY 327 LysArgTyrArgAspSerIleGlyGly-----IleGluVal--- 338
Db 2938 TTAACAGCAAGTACTTTCAGTTGGAAATGGGAATAAAGCAGTGCATCATTTGAAGTATAC 2997
QY 339 -----LysGlyIleLysAsnGluLeuThrArgGlu---IleTyrThr 351
Db 2998 ACAGATCAAGACATACCTGGAAGGGTTTGTGGAAACCTGACTTACGAATCCATTTCGTCA 3057
QY 352 ThrGluIleAsnPheAspArgLeuProGlnLeuArg-----363
Db 3058 ACTGCATATAATGTAAAGCTGGTCCACCGCTCAACCAAAAGGTCTAGTCTTCTACTAT 3117
QY 364 -----ValGlnProAsnLeuAlaThrMetGlu 372
Db 3118 GTTTCACGTATCTTACAGACAGACTCTCCGCCATGTGAGACCACCTCTTGTTCATATAGAG 3177
QY 373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyr 392
Db 3178 AGAAGCATA-----TATTTTGTATAATCTGGAAAAATACACTGATTAT 3219
QY 393 -----ThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIleSerAsnArg 408
Db 3220 ATATTAATAATTTACTCCATCAACAGAAAGGA-----3252
QY 409 AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySer 428
Db 3253 -----TTCTCTGATACCTATATCTGCCCGAGCTATAC-----3282
QY 429 ProThrThrLysThrIleArgProPheGluSerTyrLysValSerIleValThrAspArg 448
Db 3283 -----ATCAAGACTGAAGAA 3297
QY 449 GlnSerProValSerProIleGlnProHisPheIleIleAsnGlnIleGluLeuTyr 468
Db 3298 GATGTCACAGAACTTCACCAATA-----ATCAACTTTT 3333
QY 469 LeuAsnGlySerSerAsnAsnThrLeu-----LysTyrSerAlaGlyGly 483
Db 3334 AAAAACTTCTCTACTCTCAGTTCTCTATCATGGGATCCCCAGTAGTAAGCAAAATGGT 3393
QY 484 SerLeuSerAsnTyrGlnAsnThrPhePheGlnPheProArgLysLysAspCysAsn 503
Db 3394 GCAATAATAAGTTATGATTTAACT-----TTACAAGGACCAAAATGAA-----3435
QY 504 LeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIle-----520
Db 3436 -----AATTATCTTCTTCTACTTCTCTGAT 3459
QY 521 -----LeuSerHisPheSerLeuPheThrTyrSerTyrValIle 533

```



```
Db 3460 AATTACATAATATTGGAGAGCTTTCACCAATTATACATTATATAGCTTTTGTGCGCGCA 3519
Qy 534 GlyLeuGlnLeuGlnLeuLeuAaspThrGlyValLeu---GlyTrpThrHisSerSerVal 552
Db 3520 AGAACTAGAAAAGAGCTTGGTCCCTCCAGTATTCTTTTTCACACAGATGAGTCAGTG 3579
Qy 553 -----AspArgTyrAsnAlaIleSerAspLysIle----- 562
Db 3580 CCGTTAGACCTCCACAAAATTTGACTTTTAATCACTGTACTTCAGACTTTGTATGGCTG 3639
Qy 563 -----IleThrMetIleProAlaIleLysGlyAsn 572
Db 3640 AAATGGAGCCCAAGTCTCTTCCAGGTGGTATTCTTAAAGTATATAGTTTAAATAATCAT 3699
Qy 573 AsnLeuAaspThrAsnSer-----LysValIleGluGlyProGlyHisThrGlyGly 589
Db 3700 GAACATGAATGACATATATATTAAGAAATATACAGGA---TTTAAACATGAAGCC 3756
Qy 590 AsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIleThrCysGluThrProAsnSer 609
Db 3757 AAACTTGTTGGACTGGAA-----CCAGTCAGC 3783
Qy 610 ThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn---GlyAlaGlyAsnThrLeu 628
Db 3784 ACC-----TACTCTATCCGTGTATCTGCGTTCACCAAGTTGGAAATGGCAATCAATTT 3837
Qy 629 ProAsnIle-----SerLeuThrIleProGlyValIle 639
Db 3838 AGTAATGTAGTAAATAATTCACAACCAAGAATCAGTCCAGATGTCGTG 3885

RESULT 2
DQ035722 13107 bp DNA linear GSS 02-JUN-2005
LOCUS Homo sapiens DNAB5 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION Genomic survey sequence.
ACCESSION DQ035722
VERSION DQ035722.1 GI:66886931
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
MAMMALIA; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 13107)
NIELSEN,R., BUSTAMANTE,C., CLARK,A.G., GLANOWSKI,S., SACKTON,T.B.,
HUBISZ,M.J., FIEDER-ALON,A., TANENBAUM,D.M., CIVELLO,D.,
WHITE,T.J., SNINSKY,J.J., ADAMS,M.D. and CARGILL,M.
A Scan for Positively Selected Genes in the Genomes of Humans and
Chimpanzees
JOURNAL (er) PLOS Biol. 3 (6), E170 (2005)
PUBMED 15869325
REFERENCE 2 (bases 1 to 13107)
NIELSEN,R., BUSTAMANTE,C., CLARK,A.G., GLANOWSKI,S., SACKTON,T.B.,
HUBISZ,M.J., FIEDER-ALON,A., TANENBAUM,D.M., CIVELLO,D.,
WHITE,T.J., SNINSKY,J.J., ADAMS,M.D. and CARGILL,M.
Direct Submission
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment. Translation starts at the beginning of
alignment.
FEATURES
source Location/Qualifiers
1. .13107
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
<1. .>13107
/genes="DNAB5"
/locus_tag="HC4910"
Gene
ORIGIN
Alignment Scores:
3460 AATTACATAATATTGGAGAGCTTTCACCAATTATACATTATATAGCTTTTGTGCGCGCA 3519
Qy 534 GlyLeuGlnLeuGlnLeuLeuAaspThrGlyValLeu---GlyTrpThrHisSerSerVal 552
Db 3520 AGAACTAGAAAAGAGCTTGGTCCCTCCAGTATTCTTTTTCACACAGATGAGTCAGTG 3579
Qy 553 -----AspArgTyrAsnAlaIleSerAspLysIle----- 562
Db 3580 CCGTTAGACCTCCACAAAATTTGACTTTTAATCACTGTACTTCAGACTTTGTATGGCTG 3639
Qy 563 -----IleThrMetIleProAlaIleLysGlyAsn 572
Db 3640 AAATGGAGCCCAAGTCTCTTCCAGGTGGTATTCTTAAAGTATATAGTTTAAATAATCAT 3699
Qy 573 AsnLeuAaspThrAsnSer-----LysValIleGluGlyProGlyHisThrGlyGly 589
Db 3700 GAACATGAATGACATATATATTAAGAAATATACAGGA---TTTAAACATGAAGCC 3756
Qy 590 AsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIleThrCysGluThrProAsnSer 609
Db 3757 AAACTTGTTGGACTGGAA-----CCAGTCAGC 3783
Qy 610 ThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn---GlyAlaGlyAsnThrLeu 628
Db 3784 ACC-----TACTCTATCCGTGTATCTGCGTTCACCAAGTTGGAAATGGCAATCAATTT 3837
Qy 629 ProAsnIle-----SerLeuThrIleProGlyValIle 639
Db 3838 AGTAATGTAGTAAATAATTCACAACCAAGAATCAGTCCAGATGTCGTG 3885

Pred. No.: 0.536 Length: 13107
Score: 124.00 Matches: 140
Percent Similarity: 35.20% Conservative: 86
Best Local Similarity: 21.81% Mismatches: 218
Query Match: 3.20% Indels: 198
DB: 11 Gaps: 35
US-10-783-417-2 (1-735) x DQ035722 (1-13107)
Qy 97 GlyProIleGlyIleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPhe 116
Db 3676 GGACCTATTGGGAATCTTATGCTCTTAACAGATATGAGCTTCTGATAGCA----- 3729
Qy 117 TrpProAlaGlyGluGlnAaspLysThr-----ValTrpThrGlnPheIle 131
Db 3730 -----AGGGAAGAGATAGACAAAGTTGATACACTGCACATATGCTTGGGAGAGCTGCTG 3783
Qy 132 ---LysMetGlyGluIlePheValAasp---ThrProLeuThrGluSerIleLysGlnLeu 149
Db 3784 GCAGTGTCTGGGAAGTCCAGAAATAATAGTCTCCTCAGCCAGCCAGTTTCAAGAAAGAG 3843
Qy 150 LysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAasp 169
Db 3844 CTTATTAGTGTGTGGAGGTATTCCTCAAGATTGTCACCACTTTTATCTGGACTATGAT 3903
Qy 170 AspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeuProProSerSerAlaLeuGln 189
Db 3904 -----TTGAATGGTCCAAATGCTGCGCTTGAAGCCCGCAGGAAGCC---AGT 3948
Qy 190 GlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAaspPheIleArgGlu 209
Db 3949 GACAGGCTTATCATGTTTTCAGAAATCAATTTGATAATATCTATCGGAAATACATCATAT 4008
Qy 210 IleProGlyPheGlnLeu-----GluThrTyrLysThrLeuLeuLeuProIle 225
Db 4009 ACTGGAGGAGAGAGCTTTTGGCTGCGCAGCTACACAGTATCCTCAGCTTCTTGAATA 4068
Qy 226 TyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGlyAlaGluLeuAlaAasp 245
Db 4069 AAGAAGCAA-----CTAAATCTTCTACAGAAATATATATCTCTG----- 4107
Qy 246 GluTrpAsnAlaAaspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAasp 265
Db 4108 -----TACAACAGTCTCATAGAA-----ACTGTAAATAGC 4137
Qy 266 TyrTyrLysLeuLeu-----LysGluAsnIleProLysTyrSerAsnTyrCys----- 281
Db 4138 TATTATGATATCTTTTGTGTAGAGTGAATATTGAAAAAATTAAACATGAACCTCTTAGAA 4197
Qy 282 ---AlaAsnThrTyrArgThrGlyLysAsnLeuArgAaspGluProAsnMetLysTyr 300
Db 4198 TTCCAGAACACAGATCTCGAAAGCTTCCCGGGCGCTTGAAGGAC-----TGG 4242
Qy 301 SerIlePheAsnAaspTyrArg----- 308
Db 4243 CAGGCTTTTGTGGACCTGAAGAAGATCATGATGATTTTTCAGCGAGTGTTCGCCGCTGCTG 4302
Qy 309 ---TyrMetThr-----IleThrValLeuAasp 316
Db 4303 GAATACATGGCCAGTAAGCCATGATGGAGGGCGCTCGGAAAGGATATACACCTTCACC 4362
Qy 317 ThrIleSer-----GlnPheSerLeuTyrAspIleLys----- 327
Db 4363 GGGCACAGTCTGGATGTGGGAATGAAAGCTTTAAGTTAAGAAATATCATGGAGGCACCT 4422
Qy 328 -----ArgTyrArgAaspSerIleGlyIleGluValLysGlyIleLys----- 342
Db 4423 CTTCTGAAATATAAAGAGGAATAGAGGACATCTGTATCAGTGTGGTGAAGAGAGAGAC 4482
Qy 343 -----AsnGlu----- 344
Db 4483 ATTGAGCAAAAGCTGAAGCAAGTGAATTAATGAGTGGGCAATATAAACATTCACCTTCGGC 4542
```

QY 345 -----LeuThrArgGluIleTyrThrThrGluIleAsnPhe 356
Db 4543 AGCTTTAAACCGTGGAGAGCTCTCTTGAGAGGAGACAGTACCTCGGAATCATCGCC 4602
QY 357 AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThr 376
Db 4603 AACATGGAGACAGCTTGATGTTGCTGGGATCCCTACTGAGCAACAGGTACATATG--- 4659
QY 377 ArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThrGluAsn--- 395
Db 4660 -----CCATTCAA-----GCCAGATTCAAAATGGGTGCGAGTACCTTTCCAACTCA 4707
QY 396 ThrAsnPheGlyAsnArgLeuValGlyIleSerAsn-----ArgAspAlaPro 411
Db 4708 ACAGACATCATCGAGAGCTGGATGACGAGCTGCAAAACCTGTGGATTTATTAGAGCTGTC 4767
QY 412 ThrGlySerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySerProThrThr 431
Db 4768 TTGTGGGAGGAGACATGCGCAAGCAGCTGCCCAAGGAA-----GCC 4809
QY 432 LysThrIleArgProPheGluSerTyrLysValSerIleValThrAspArgGlnSerPro 451
Db 4810 AAGCGTTTCTTAACATAGATAAATCTGGTGAAGATCATGACTCGGCGACATGAAGTG 4869
QY 452 Pro-----ValSerProIleGlnProHisPhe 460
Db 4870 CCCAGTGTAGTCAGTGTGTGTGGAGATGAGACCTCGGGCAGCTGTGTACACAC--- 4926
QY 461 IleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsnThrLeuLysTyrSer 480
Db 4927 TTGCTGGACAGTGGAAATA----- 4947
QY 481 AlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPheGlnPheProArgLysLys 500
Db 4948 TGCCAGAAATCCCTTACTGGTACTTGGAGAAAACAGACTGCTTCTCTCGG----- 5001
QY 501 AspCysAsnLeuValIleAspProGly-----CysSerProAsn 513
Db 5002 ---TTTTCTTCTCTCAGATCTCGCCCTTCTAGAGATTCTGGGGCAGGGCTCGGACTCC 5058
QY 514 PheAsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIle 533
Db 5059 CACACTATACAGCCCATTTGCTGAATGTGTTGAC----- 5094
QY 534 GlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTrpThrHisSerValAsp 553
Db 5095 -----AACATTAATCTGTCAAG 5112
QY 554 ArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsn 573
Db 5113 TTCCAGAAAGATCTATGATCGAATCTGTCA--ATTCTCTCAAGAGGGTGAGACG 5169
QY 574 LeuAspThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeu---Val 592
Db 5170 ATTGAA-----TTGGATAAACCTGTCTATGCGAGAGGCAATGTGAAGTT 5214
QY 593 TyrLeuGlnSerGln--GlyArgLeuGluIle-----ThrCysGluThrProAsn 609
Db 5215 TGCTTAAATCTCTTTTGAAGAATCTCAGTCTCTCATTTGCTGATTCGCCAGGCA 5274
QY 609 eThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuP 629
Db 5275 GCGGCAAA-----TATTCAAGAAACAGGTTTCACTAACTGAATCTTTTCATCTTTC 5328
QY 629 ro 629
Db 5329 CC 5330

RESULT 3

LOCUS CL977084 3477 bp DNA linear GSS 21-SEP-2004
DEFINITION OsifCC029775 Oryza sativa Expressed Library Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CL977084
CL977084.1 GI:52408678
GSS.

Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)

Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

1 (bases 1 to 3477)

AUTHORS

Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G.K.S., Deng, X.W. and Wang, J.

TITLE

An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis

JOURNAL

Unpublished (2004)

COMMENT

Contact: Chen Chen
Department of Bioinformatics
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.

FEATURES

source

1. 3477
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Expressed Library"
/note="Oryza sativa exon trapped genomic sequences "

ORIGIN

Alignment Scores:
Pred. No.: 0.122 Length: 3477
Score: 121.50 Matches: 144
Percent Similarity: 31.81% Conservative: 79
Best Local Similarity: 20.54% Mismatches: 243
Query Match: 3.14% Indels: 235
DB: 10 Gaps: 31

US-10-783-417-2 (1-735) x CL977084 (1-3477)

QY 132 LysMetGlyGluIlePheValAsp---ThrProLeuThrGluSerIleLys----- 147
Db 1105 AAAATGAACAACTTACCTAGACCAAAACCAATACGGGTTCCATCCAAAAGAAATA 1164
QY 148 -----GlnLeuLysLeuGlnThrLeuGluGlyPheArg 158
Db 1165 GGAATGCTAGCAACCTGCAAGTTTTCAGCTAAGCAACACACCTTGAGCGG----- 1218
QY 159 GlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGln 178
Db 1219 -----GAGATACCCACAGCTCTAGCAAAATCTGACCAACCTAGTACTTTGAAA 1266
QY 179 AlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaLeuLeuLeu---LysIle 197
Db 1267 CTATATGTAATGACTGTGGGCGCTATACCCAAAACCTCTGCACCTCACCAGATG 1326
QY 198 ArgPheGluAsnVal---HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGlu 216
Db 1327 CAATTACTTAGTCTTAGCAAAAACAAACTTACTCGGAAATCCCTCGTGTATTATCAAT 1386
QY 217 ThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeu 236
Db 1387 CTAACAAAAGTGGAAAAAATTTTACTATACCAAAACCAA----- 1425
QY 237 LeuGlnGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIle 256
Db 1425 ----- 1425
QY 257 GluProAsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLys 276
Db ---:|||||

```

1426 Db -----GTCACGGGTTCATCCAAA 1446
277 QY TyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro 296
1447 Db GAAATAGGAATGCTCCCAACCTGCACGCTTTTAGGTCTAGGCAACAATACCTTGAACGGA 1506
297 QY AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAsp 316
1507 Db GAGATACCCACAACCTCTATCAAT-----CTAACCAACCTAGAT 1545
317 QY ThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyIle 336
1546 Db ACCTTG-----TCCCTATGGAT----- 1563
337 QY GluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPhe 356
1564 Db -----AATGAACCTTCGGGGCATATACCCCAAAACCTGCACGCTC 1605
357 QY AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyr----- 373
1606 Db ACCAGATGCAATACCTTAGTCTTAGCAGCAACAACCTTACTGGCGAAATCCCTGGGTGT 1665
374 QY -----AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlu----- 387
1666 Db TTATCCAATCTACAAAATGGAA--AACTTTACCTATACCAAAACCAAGTCACGGGT 1722
388 QY -----GlnPheIlePheTyrThrGluAsn 395
1723 Db TCCATCCAAAAGAAATAGGAATGCTTCCCAACCTGCAGAGTTTTCAGAGTAAAGCAACAAC 1782
396 QY ThrAsnPheGlyAsnArgLeuValGlyIleSerAsnArgAlaProThrTyrSerAsn 415
1783 Db ACCTTGACGGGGAGATATCCACAGCTCTATCAAT-----CTAACCAAC 1827
416 QY ThrIleThrGluThrLeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArg 435
1828 Db CTAGCTATTTCCTCCCTATGGGTAAAGAACTTCGGGGCTATACCCCA----- 1878
436 QY ProPheGluSerTyrLysValSerIleValThrAspArgGlnSerProProValSerPro 455
1879 Db -----AAACTCTGCATCTCACCAGATCCAAATATCTTGATCTTAGCAGT 1923
456 QY IleGlnProHisPheIleIleAsnGln-----IleGluLeuTyrLeuAsnGlySer 472
1924 Db TCTCTTCTCGGAAATTCGAAAATCTCAGCGGCATAGCAGACCTTTGGCTTGATAATAAC 1983
473 QY Ser-----AsnAsnThrLeuLysTyrSerAlaGlyGlySerLeuSerAsnTyrGln 489
1984 Db TCATTTTCAGGACATTTGCCTGCAAAATGTATGATGGTGGGACACTCAAAAACCTTTCATG 2043
490 QY -----AsnThrThrPhePheGlnPheProArgLys---LysAspCys----- 502
2044 Db ATTGGTGGGAATGCCTTCGATGGCCCCATTCGGAAGAGCTTGAAGACATGTACAGCTTG 2103
503 QY -----AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsn 516
2104 Db GTTAAACCTTTCGCGTACAATAACCTGCTTACGGGACACATATCTGAACATTTTGGAGTG 2163
517 QY TyrSerHisIleLeuSer-----HisPheSerLeuPheThrTyr 529
2164 Db TATCCGCATCTCAAAATCGGTGAGCTTGCATACAAATAGGTTCTTTGGCGACATCTCACA 2223
530 QY SerTyrValIleGlyLeuGlnLeuIleLeuAsp-----ThrGly 543
2224 Db AATGGTGTCTAGTCCCGAGCTTGAGAAATGATTTTCACAAAACATGATCAGAGT 2283
544 QY ValLeuGlyTrpThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIle 563
2284 Db TTACTCAGACTTGATCAT-----AATAATATCATGTGGTGAG----- 2319
564 QY ThrMetIleProAlaIleLysGlyAsn-----AsnLeuAspThr 576
2320 Db -----ATTCAGCAGAAATTTGGAATCTGAAAAGCCTATATAGATAAAATTTATCATTC 2373

```

577 AsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSer 596
2374 AAT-----CAGCTCTCAGGATATTATACAGCAGCAGCAGTAAAGTTAAGCAAT 2421
597 GlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616
2422 TTAGGTACCTTGATGATCAAGAAC----- 2448
617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
2449 -----AATTGAGTGGACCAATATCCC 2469
637 GlyValIleGly-----IleProProGln 644
2470 GATGAACCTAGGGGACTGCATCAGTTACAGGGAAGTCTTCCATCATTCGTCAGCGCG 2529
645 ArgLeuAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGly 664
2530 ATGTGAGCTTAAACAGTTCTTGTGTGTCTTCAACAACTTGGAA----- 2574
665 TyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIlePro---PheIle 683
2575 ---GGACCTCTCCAGCAGGACATCTACTTCAA---AATGCATCAATAAGCTGTTATT 2628
684 PheAsnArgAlaAspValSerAsnSer----- 692
2629 CACAATAAGTTATAGCATCTGGTCAACCAAGCAGTACTAAGTTTGTACTGCCCA 2688
693 -----IleLeuIleIleAspLysIleGluPheIleProIleThrSerSerMet 708
2689 ATTGTTCTTGTGTGGTCATTGTCTATTCGCTACAAATATTGTATCATCAAGCTGTT 2748
709 HisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePhe 728
2749 CATTAACAGAGAAAACAAACAGCAAGCTAGTGTCTATTCTGTCAGGAAGCATGTTT 2808
729 Thr 729
2809 TCT 2811

RESULT 4
AG392993/c 1754 bp DNA linear GSS 21-DEC-2004
LOCUS Mus musculus molossinus DNA, clone:MSMg01-211C07.TJ, genomic survey
DEFINITION sequence.
ACCESSION AG392993
VERSION AG392993.1 GI:48018548
KEYWORDS GSS.
SOURCE Mus musculus molossinus (Japanese wild mouse)
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
AUTHORS Abe, K., Noguchi, H., Tagawa, K., Yuzuriha, M., Toyoda, A., Kojima, T.,
Ezawa, K., Saitou, N., Hattori, M., Sakaki, Y., Moriwaki, K. and
Shiroishi, T.
TITLE Contribution of Asian mouse subspecies Mus musculus molossinus to
genomic constitution of strain C57BL/6J, as defined by BAC-end
sequence-SNP analysis
Genome Res. 14 (12), 2439-2447 (2004)
15574823
JOURNAL
PUBMED 2 (bases 1 to 1754)
REFERENCE Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
AUTHORS Direct Submission
TITLE
JOURNAL

COMMENT
Library availability, please contact Kuniya Abe (abe@tc.riken.jp).

Db	1041	-----GGCCTTATATCCAGGTCTATGACCAACGCCCATATG---	1076
Qy	547	TrpThrHisSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIle	566
Db	1076	-----	1076
Qy	567	ProAlaIleLysGlyAsnAsn-----LeuAspThrAsnSerLysValIleGlu	582
Db	1077	-----TGCCAAGGTAAATGAGCAGCAGCCATATATGCCAGGTAATGAGCAGCGCCATAT	1130
Qy	583	GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle	602
Db	1131	ATCCAGGTCTATGAGCAGCGGTCTATATGCCAGCTCAATCTCAGTCTCAGTCTCAACCT	1190
Qy	603	ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn	622
Db	1191	CAACCTCAA---CCTCAACTCAGCAACATATATGCCTGCAGCGCAACCTCGAATCAAC	1247
Qy	623	GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIle-----	639
Db	1248	ATCAGCGGCCCTCTTCAACCTTAACCAAGTATCTACCAACCACCGGGTTAGTTCTTAGCCCA	1307
Qy	640	-----GlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThr-----	653
Db	1308	GTGCCTCATATCTCTCTCTCTCGGTCAATGGCTGTACTGTGACGCCTACTCTCAG	1367
Qy	654	-----AsnTyrAsnAsnLeuGlnTyrGlyAspPhe	663
Db	1368	CCCTCTCCCATGTTTCAGTCAGATGATGGGATTTCTTCTCTCGATATACGGTTTT	1427
Qy	664	GlyTyrPheGlnPheProSerThrValThrLeu---ProLeuAsnArgAsnIleProPhe	682
Db	1428	GGACCACTGCAGTCACCTACATCCCAATTCTTCAACCATCTCTACCGGTTACCCGAAT	1487
Qy	683	IlePheAsn-ArgAlaAspValSerAsn-----SerIleLeuIleI	696
Db	1488	ATGTTCTCCCGAGTCGCTTACCCATTGCTGTACCAGAGTTCAGTATCTCTC-----	1542
Qy	696	eAspLysIleGluPheIleProfileThrSerSerMetHisGlnAsnArgGluLysGlnIly	716
Db	1543	-----AACCATTACCCCAACTTCTCATTTTCACAAATTGCTCAACAGGAAGT	1592
Qy	716	sLeuGluThrIleGlnThrLysIleAsnThrPhePheThr	729
Db	1593	CTTGGACCGGTGCAGTCCAAGTCAAGGTCTCTCTCTCAGC	1632
RESULT 6			
LOCUS	CL967047	6777 bp	DNA linear
DEFINITION	CL967047.1	GI:52388742	GSS
ACCESSION	CL967047		
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			

```

Rice genomic sequence.
Class: exon-trapped.
FEATURES
    source              Location/Qualifiers
    1..6777
    /organism="Oryza sativa (indica cultivar-group)"
    /mol_type="genomic DNA"
    /db_xref="taxon:39946"
    /clone_lib="Oryza sativa Express Library"
    /note="Oryza sativa exon trapped genomic sequences "
ORIGIN
Alignment Scores:
Pred. No.:      0.615      Length:      6777
Score:          119.50     Matches:    145
Percent Similarity: 34.07%  Conservative: 85
Best Local Similarity: 21.48%  Mismatches: 209
Query Match:      3.09%     Indels:    238
DB:               10       Gaps:      36

US-10-783-417-2 (1-735) x CL967047 (1-6777)
QY 151 LeuGlnThrLeuGluGlyPheArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAsp 170
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 427 ATATCATCGTTACTAGATTTCAGGCAAGCCATCAATGATCAACGGGAGCGCTAAACAGT 486
QY 171 Trp-----
   |||
Db 487 TGGAGCAGCGCGGTTCCCAATGCCAGTGGAGGGGCTCAGTTGTAGCCGGAGCGCAGTC 546
QY 175 LysArgLeuGlnAla-----ProGlyLeuProProSerSerAlaLeuGln 189
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 547 GCGCGAGTCAGCGCGCTCAACCTTACCAGAAAGTTTATCGGGTCCATTTCTGCCTCT 606
QY 190 GlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPheIleArgGlu 209
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 607 GTTGGAACTGCATTCCTTCATACACTTGACTGTCCCATTAATTAACCTGTGCGGGCAG 666
QY 210 IleProGlyPheGlnLeuGluThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAla 229
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 667 ATGCCT-----
QY 230 AlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTrpAsnAla 249
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 673 -----CATCTTAACAATCTGCAGAGATGACAGGTCCTTAATCTGAGCTACAATTC 723
QY 250 -----AspIleHisPro-----
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 724 TTAGATGAATTAATCCAGACAGCTCACCATTGTTCCAACTACAGCAATTAATCTTA 783
QY 254 -----SerGlnIleGluProAsnAlaGly-----ThrSerAspAsp 265
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 784 AGATATAATAAATTCGGAAGTCGTAATCTCTCGCAATAGGTCTCTTACCAACCTTGT 843
QY 266 TyrTyrLysLeuLeuLysGluAsnIle-----ProLysTyrSerAsnTyr 280
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 844 TACATGTCCTATGCCAATAATATCTCTCTGCAATTATTCCCCCAACCCCTCGGTAACATG 903
QY 281 CysAlaAsnThrTyrArgThrGlyLeuLysAsn-----LeuArgAspGlu 295
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 904 ACCTCGTAGTAATCTCGCTCTTACACMAATCATCTTAGAGGAAGAGTCCCGATGAA 963
QY 296 -----ProAsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThr 311
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 964 ATTGGGAACCTTCCCAATATCAGCATCGTAGTCTCTAGAGAT-----AACATGTTGTCA 1017
QY 312 IleThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAsp 331
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1018 GGTAGTATACCAAGAACCCCTTCT-----AATCTCTTACGGGG-----AATCT 1047
QY 332 SerIleGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThr 351
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1048 TCTATTACAGGAATAAGATTACGTGCA-----AATCTCTTACGGGG-----ACT 1092

```

```

352 ThrGluIleAsnPhe---AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThr 370
1093 CTACCTATAAACATGGGTGATAATCTCCCTAAATCTCAAA-----TCTCTCTTC 1140
QY 371 MetGluTyrAsnLeuThrArg-----AlaSerPheLysLeuPheSerPheLeu 386
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1141 TTGGGCCCAACAATGCTTGGGGTCACATCCAGCTTCTTAGGCAACGCTTCAACGCTA 1200
QY 387 GluGlnPheIlePheTyrThrGluAsn-----ThrAsnPheGlyAsn 400
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1201 AAACAAGTAGTTCTACAACAACCAATTCACCTGGCCAGATTCCTACTACCTTTGGA--- 1257
QY 401 ArgLeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1258 AAGCTTTTGGACTGTCCAGACTAGATCTTCAGGCA-----AACATGCTTGAAGCA 1308
QY 421 LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1309 AAAGACAGTGAAGGCTGGGA-----TTTCTGTATGCATCGAGAACTACAGTGTCTTA 1362
QY 441 LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1363 CAAATCTTCAATTAACATGAGAACTCACTACAGGGAATTATACCA----- 1407
QY 461 IleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsnThrLeuLysTyr--- 479
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1408 -----AATTCAAATC-----GGTAGCCTATCGAACACTCTTTGGGTATCTC 1446
QY 480 ---SerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArg 498
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1447 ATGTGGGTGGAATGAATCATCTATCTAGC----- 1473
QY 499 LysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSer 518
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1474 -----ATAGTTCCCTCATGATAGGGAACCTTCTGGGTAGTT 1512
QY 519 HisIleLeuSerHisPheSerLeuPheThrTyrSerTyr-ValIleGlyLeuGln----- 536
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1513 CAACTGCATTAGATGCAACAAGCAAGGTTTCGATCGATCCATGATGATGGAATCTT 1572
QY 537 -----LeuGlnIleLeuAsp---ThrGlyValLeuGlyTrpThrHisSerSe 551
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1573 AATACTTGGAAATATTATACATCTTGGAAAGAACAGGTTCACTGGGCTTATTCATAT-TC 1631
QY 551 rValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysG 571
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1632 TATTGGCAGACTTACACGGCTGACAGAG----- 1659
QY 571 YAsnAsnLeuAspThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLe 591
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1659 ----- 1659
QY 591 uValTyrLeuGlnSer-----GlnGlyArgLeuGluIleThrCysGluThrProAs 608
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1660 -CTCTATCTGGAAGAACCAATGCATTGGAAGAACCATTCCTCAATTTCT- 1704
QY 608 nSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLe 628
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1705 -----CTTGGCAACCTCAAGGCTTGAGCGT 1730
QY 628 uProAsnIleSerLeu-----ThrIleProGlyValIleGlyIleProPr 643
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1731 ACTAAATCTTTCCATAACATCTTATCGGGCACCATCCAGCAGTACTGGGGGATCTACC 1790
QY 643 o-----GlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGl 661
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1791 TCTACTCAGCAAGCTGGAC-----CTTTCTTACATAATATCTCCAA---GG 1832
QY 661 YAsp-----PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnAr 678
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1833 AGAAATACCAGGATGAATTAATTCAGAAATGCTACATCTGTTTATCTTTGAGGCAATAG 1892
QY 678 gAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIleLeuIleAspLys 698

```



```

Db      1893  GGGACTATG-----CGGAGGACTTATGGATCTCCATATGCCATC----- 1931
Qy      698  sileGluPheIleProIleThrSerMetHisGlnAsnArgGluLysGlnLysLeuG1 718
Db      1932  ::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :
Qy      718  uThrIleGlnThrLysIleAsnThrPhePheThrAsnHisThr 732
Db      1976  CGAGATTGTGATCCCATAGTTGGTTCTTCTGTCACTCACAGT 2018

RESULT 7
LOCUS   AY416726
DEFINITION Mus musculus LY64 gene, VIRTUAL TRANSCRIPT, partial sequence,
GENOMIC SURVEY SEQUENCE.
ACCESSION AY416726
VERSION   AY416726.1 GI:39772686
KEYWORDS GSS.
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1986)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrier,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1986)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrier,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1. .1986
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1986
/gene="LY64"
/locus_tag="HCM5986"
gene

ORIGIN
Alignment Scores:
Pred. No.: 0.143 Length: 1986
Score: 117.50 Matches: 114
Percent Similarity: 31.72% Conservative: 76
Best Local Similarity: 19.03% Mismatches: 190
Query Match: 3.04% Indels: 221
DB: 10 Gaps: 22

US-10-783-417-2 (1-735) x AY416726 (1-1986)
Qy      201  AsnValHisAsnAspPheIleArgGluIleProGly----- 212
Db      112  AACTGTGAATAATTAGTCTCAATGAATAATCTGGCACTTTACCAACTCAACAGAAATGT 171
Qy      213  -----PheGlnLeuGluThrTyrlsThrLeuLeuLeuProIleTyrlsAlaGlnAla 230
Db      172  TTGGAGTTGAGCTTTAATGTCTTGCCTACCATTCATAAACACAGCCTTCAGAGACTTATC 231
Qy      231  AsnPheHisLeuAsnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 246

```

```

232  AATCT-CACCTTCCTGGATTTAACCCAGGTGCCAGATTTTACTGGATACATGAAGATACTTT 290
246  ----- 246
291  CCAAGGCCAACATCGGTGTAGACACACACTTTGTACTAACCGCAATCCCTGATATTTATGGC 350
247  -----TTPAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGly 261
351  AGACACAGCAGCTTAGTGGGCTTAAGGCAATGAAGCATCTGTCTTCATCCAAACAGGAAT 410
262  -ThSerAspAspTyrlsLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 281
411  ATCCAGTATTGATTTATCCCACTGCACAAATCAAAAACCTTGGAA--AGTCTCTATCT 467
281  sAlaAsnThrTyrlsArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSe 301
468  CGGAAGCAACCATATTTCTCCATTAAGCTCCCAAGGTTTCCCAACAGAGAAGCTGAA 527
301  rIlePhe-----AsnAspTyrlsArgTyrlsMetThrIleThrValLeuAspThrI1 318
528  GGTCTGTGATTTTCAGATAATATGCTATCCATTACCTGTCTAAAGAGATATGAGCTCTCT 587
318  eSerGlnPheSerLeuTyrlsAspIleLysArgTyrlsArgAspSerIleGlyIleGlu-- 337
588  CCACAGCCCACTAATCTGAGCCTTAACCTTAATGGAATGACATTCAGAGATAGAGCC 647
337  ----- 337
648  GGGGGCTTTTCAGCTCAGCTGCTTCCAAAGTTTAAACTTTTGGAGGAGCATCAGAACTTGCT 707
338  -----ValysGlyIleLysAsnGluLeuThrArgGluIleTyrlsThrGluIleAs 355
708  AGTTATCTTCAAGGGTTTGAAGAACTCTACGATCCAGTCTCTTTGGCTGGGACATTTGA 767
355  n-----PheAspArgLeuProGlnLeuArgVa 364
768  GGACATGATGACGAAGATATTAGTCCTGCCGTGTTTGAGGGTCTCTGTGAATATGCTGT 827
364  lGlnProAsnLeuAlaThrMetGluTyrls-----AsnLeuThrArgAlaSerPheLysLe 382
828  GGAGAGATCAACCTACAGAAGCATTTCTTCAACATTTCTCCCAACACACATTCCTATG 887
382  uPheSerPheLeuGluGlnPhe---IlePheTyrlsThrGluAsnThrAsnPheGlyAsnAr 401
888  CTTTCAGTGGCTCCAGGAACCTGAGCTTAACAGCCACTCAGCTGAGTGAATTCGCTTCTGG 947
401  gLeuValGlyIleSerAsnArgAspAlaProThrTyrlsSerAsnThrIleThrGluThrLe 421
948  ACTTGTGGGACTAAGC----- 963
421  uTyrlsGluArgThrGlySerProThrTyrlsThrIleArgProPheGlu-SerTyrl 441
964  -----ACACTTAAGAAATAGTTCTCAGTGC 989
441  ysValSerIle---ValThrAspArgGlnSerProValSerProIleGlnProHisP 460
990  AATAAGTTTGAAGAAATTTGTGCCAAATCAGTGTCTTCAATTTCCGCTCC-CTTACTCACC 1048
460  lleIleAsnGlnIleGluLeuTyrlsLeuAsnGlySerSerAsnAsnThrLeuLysTyrls 480
1049  TTTCCATC-----AAGGGCAACAGAGAGACTTG 1078
480  erAlaGly---GlySerLeuSerAsnTyrlsGlnAsnThrThrPhePheGlnPhePro---- 497
1079  AGTCGGTACTGGCTGTGTTAGAAAACCTAGAAAATCTCCGCAACTTGACCTCAGCCATG 1138
498  -----ArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheA 515
1139  ATGACATTGAACCTCTGATTGT-----TGCAACCTACAACTCC 1177
515  snAsnTyrlsHisIleLeuSer-----HisPheSerLeuPheT 528
1178  GAAACCTGTCTCACTTACAGAGCCTGAACCTTGAAGCTACAATGAACCCCTGAGCCTCAAGA 1237

```



```
QY 528 hrTy-SerTyrrValIleGlyLeuGlnLeuGlnIleLeuAspThrGly----- 543
Db 1238 CTGAGCATTCAGAAATCCCTCAGCTAGACTCTAGATTTGGCATTTACTCGACTAA 1297
QY 544 -----ValLeuGlyTrpT 548
Db 1298 AGTGAAAGATGCACAGAGTCCCTCCAGAACCTCCATCTTTGAAGGTGCTGAAATCTT 1357
QY 548 hrHisSerValAspArgTyrAsnAlaIleSerAspIlyIleIleThrMetIleProA 568
Db 1358 CCCACAGCTCTTGAC-----ATCAGCAGTGCAGAGCTCTTCGATGGCTGCCAG 1408
QY 568 laile-----LysGlyAsnAsnLeuAspThrA 577
Db 1409 CACTCCAGCATTTGAACCTTACAGGGAATCACTTTCAAAAGGGAATATCCAAAAGACCA 1468
QY 577 snSerIysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerG 597
Db 1469 ACTCA-----CTTCAGACAC 1483
QY 597 lngIyArgLeuGluIleThr-----CysGluThrProAsnSerThrGlnSert 613
Db 1484 TGGGAGACTAGAAATCTGGTTTATCTCTTTTGATCTCTCTCCATGACACGACAG 1543
QY 613 yrPheIleArgLeuArgTyrAlaThr----- 621
Db 1544 CCTTCACCATCTGAAGATGATGAATCATGTAGACCTGAGTCAACAAGGCTGCATCCA 1603
QY 622 -----AnGlyAlaGlyAsnThrL 628
Db 1604 GTAGCATTCAGGCTCTTACTCATCTTAAAGGGATCTTACTCTCGGCTCCAAAT--- 1659
QY 628 euProAsnIleSerLeuThrIleProGlyValIleGlyIleProProGlnArg 645
Db 1660 -----CACATGATCATCTTACCAGTCTCTCCCATCTTGTCACGAG 1707

RESULT 8
AK089255
LOCUS
DEFINITION
Mus musculus NOD-derived CD11c +ve dendritic cells cDNA, RIKEN
full-length enriched library, clone:F630107B15 product:lymphocyte
antigen 78, full insert sequence.
ACCESSION
AK089255
VERSION
G1:26354375
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
AUTHORS
Carninci,P. and Hayashizaki,Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
PUBMED
10349636
REFERENCE
2
AUTHORS
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)
PUBMED
11042159
REFERENCE
3
AUTHORS
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,U.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
```

JOURNAL
PUBMED
REFERENCE
AUTHORS

Genome Res. 10 (11), 1757-1771 (2000)
11076861

TITLE
JOURNAL
REFERENCE
AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

TITLE
JOURNAL
REFERENCE
AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

TITLE
JOURNAL
REFERENCE
AUTHORS

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Hayatsu,N., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akashira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0445, Japan (E-mail: genome-res@gsr.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

TITLE
JOURNAL

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
Trust/MRC building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/
Location/Qualifiers
1. 2157
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM DB:F630107B15"
/db_xref="taxon:10090"
/clone="F630107B15"
/cell_type="NOD-derived CD11c +ve dendritic cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
134. 2119
/note="unnamed protein product; lymphocyte antigen 78
(MGD|MGRI:1194924, GB|NM_008533, evidence: BLASTN, 99%,
match=2136)
putative"
/codon_start=1
/protein_id="BAC40816.1"
/db_xref="GI:26354376"
/translation="MAPDIDSCFFLVALFLASCRATTSDDQKICKEVKNKYNCENLGL
NEIPGLTPNTECLFSEFNVLPTQNTFFRLNLITFDLRCQIYWIHEDTFQSOHR
LDLVLVTANPLIFMAETLSPKALHFLFIQTGISIDFPLNKLNSLGLGSNH
ISSIKLPKGPTEKLVLDQFNAAHYLSKEDSSLOQATNLNLNGNDIAGIELGA
EDSAVQSLNFGGTQNLVIFKGLKNSTIQSLWLTGTFEDMDDEIDSPAVFGLCEMSV
FESNLQKHYFNITSSNTHCFGLQELDLTSLHSELPSGLVGLTKLVLANKFE
NLQCIKASNPFSPLTHLSIKNTKRLGLEGTGLENLENRLDLSDHDDTDCNQLQ
RNLSHLSQSLNSLYNEPLSKTEAPKCEPQLLELLDAPTRLVKQKQSPFNHLHLKVL
NLSHSLDIDSSEQLFDGLPALQHLNQGHPKGNIOKTNLSQTLGLRLEILVLSFCDL

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
Trust/MRC building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/
Location/Qualifiers
1. 2157
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM DB:F630107B15"
/db_xref="taxon:10090"
/clone="F630107B15"
/cell_type="NOD-derived CD11c +ve dendritic cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
134. 2119
/note="unnamed protein product; lymphocyte antigen 78
(MGD|MGRI:1194924, GB|NM_008533, evidence: BLASTN, 99%,
match=2136)
putative"
/codon_start=1
/protein_id="BAC40816.1"
/db_xref="GI:26354376"
/translation="MAPDIDSCFFLVALFLASCRATTSDDQKICKEVKNKYNCENLGL
NEIPGLTPNTECLFSEFNVLPTQNTFFRLNLITFDLRCQIYWIHEDTFQSOHR
LDLVLVTANPLIFMAETLSPKALHFLFIQTGISIDFPLNKLNSLGLGSNH
ISSIKLPKGPTEKLVLDQFNAAHYLSKEDSSLOQATNLNLNGNDIAGIELGA
EDSAVQSLNFGGTQNLVIFKGLKNSTIQSLWLTGTFEDMDDEIDSPAVFGLCEMSV
FESNLQKHYFNITSSNTHCFGLQELDLTSLHSELPSGLVGLTKLVLANKFE
NLQCIKASNPFSPLTHLSIKNTKRLGLEGTGLENLENRLDLSDHDDTDCNQLQ
RNLSHLSQSLNSLYNEPLSKTEAPKCEPQLLELLDAPTRLVKQKQSPFNHLHLKVL
NLSHSLDIDSSEQLFDGLPALQHLNQGHPKGNIOKTNLSQTLGLRLEILVLSFCDL

FEATURES /
source

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
Trust/MRC building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/
Location/Qualifiers
1. 2157
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM DB:F630107B15"
/db_xref="taxon:10090"
/clone="F630107B15"
/cell_type="NOD-derived CD11c +ve dendritic cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
134. 2119
/note="unnamed protein product; lymphocyte antigen 78
(MGD|MGRI:1194924, GB|NM_008533, evidence: BLASTN, 99%,
match=2136)
putative"
/codon_start=1
/protein_id="BAC40816.1"
/db_xref="GI:26354376"
/translation="MAPDIDSCFFLVALFLASCRATTSDDQKICKEVKNKYNCENLGL
NEIPGLTPNTECLFSEFNVLPTQNTFFRLNLITFDLRCQIYWIHEDTFQSOHR
LDLVLVTANPLIFMAETLSPKALHFLFIQTGISIDFPLNKLNSLGLGSNH
ISSIKLPKGPTEKLVLDQFNAAHYLSKEDSSLOQATNLNLNGNDIAGIELGA
EDSAVQSLNFGGTQNLVIFKGLKNSTIQSLWLTGTFEDMDDEIDSPAVFGLCEMSV
FESNLQKHYFNITSSNTHCFGLQELDLTSLHSELPSGLVGLTKLVLANKFE
NLQCIKASNPFSPLTHLSIKNTKRLGLEGTGLENLENRLDLSDHDDTDCNQLQ
RNLSHLSQSLNSLYNEPLSKTEAPKCEPQLLELLDAPTRLVKQKQSPFNHLHLKVL
NLSHSLDIDSSEQLFDGLPALQHLNQGHPKGNIOKTNLSQTLGLRLEILVLSFCDL

FEATURES /
source

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
Trust/MRC building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/
Location/Qualifiers
1. 2157
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM DB:F630107B15"
/db_xref="taxon:10090"
/clone="F630107B15"
/cell_type="NOD-derived CD11c +ve dendritic cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
134. 2119
/note="unnamed protein product; lymphocyte antigen 78
(MGD|MGRI:1194924, GB|NM_008533, evidence: BLASTN, 99%,
match=2136)
putative"
/codon_start=1
/protein_id="BAC40816.1"
/db_xref="GI:26354376"
/translation="MAPDIDSCFFLVALFLASCRATTSDDQKICKEVKNKYNCENLGL
NEIPGLTPNTECLFSEFNVLPTQNTFFRLNLITFDLRCQIYWIHEDTFQSOHR
LDLVLVTANPLIFMAETLSPKALHFLFIQTGISIDFPLNKLNSLGLGSNH
ISSIKLPKGPTEKLVLDQFNAAHYLSKEDSSLOQATNLNLNGNDIAGIELGA
EDSAVQSLNFGGTQNLVIFKGLKNSTIQSLWLTGTFEDMDDEIDSPAVFGLCEMSV
FESNLQKHYFNITSSNTHCFGLQELDLTSLHSELPSGLVGLTKLVLANKFE
NLQCIKASNPFSPLTHLSIKNTKRLGLEGTGLENLENRLDLSDHDDTDCNQLQ
RNLSHLSQSLNSLYNEPLSKTEAPKCEPQLLELLDAPTRLVKQKQSPFNHLHLKVL
NLSHSLDIDSSEQLFDGLPALQHLNQGHPKGNIOKTNLSQTLGLRLEILVLSFCDL

FEATURES /
source

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
Trust/MRC building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/
Location/Qualifiers
1. 2157
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM DB:F630107B15"
/db_xref="taxon:10090"
/clone="F630107B15"
/cell_type="NOD-derived CD11c +ve dendritic cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
134. 2119
/note="unnamed protein product; lymphocyte antigen 78
(MGD|MGRI:1194924, GB|NM_008533, evidence: BLASTN, 99%,
match=2136)
putative"
/codon_start=1
/protein_id="BAC40816.1"
/db_xref="GI:26354376"
/translation="MAPDIDSCFFLVALFLASCRATTSDDQKICKEVKNKYNCENLGL
NEIPGLTPNTECLFSEFNVLPTQNTFFRLNLITFDLRCQIYWIHEDTFQSOHR
LDLVLVTANPLIFMAETLSPKALHFLFIQTGISIDFPLNKLNSLGLGSNH
ISSIKLPKGPTEKLVLDQFNAAHYLSKEDSSLOQATNLNLNGNDIAGIELGA
EDSAVQSLNFGGTQNLVIFKGLKNSTIQSLWLTGTFEDMDDEIDSPAVFGLCEMSV
FESNLQKHYFNITSSNTHCFGLQELDLTSLHSELPSGLVGLTKLVLANKFE
NLQCIKASNPFSPLTHLSIKNTKRLGLEGTGLENLENRLDLSDHDDTDCNQLQ
RNLSHLSQSLNSLYNEPLSKTEAPKCEPQLLELLDAPTRLVKQKQSPFNHLHLKVL
NLSHSLDIDSSEQLFDGLPALQHLNQGHPKGNIOKTNLSQTLGLRLEILVLSFCDL

FEATURES /
source

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
Trust/MRC building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/
Location/Qualifiers
1. 2157
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM DB:F630107B15"
/db_xref="taxon:10090"
/clone="F630107B15"
/cell_type="NOD-derived CD11c +ve dendritic cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
134. 2119
/note="unnamed protein product; lymphocyte antigen 78
(MGD|MGRI:1194924, GB|NM_008533, evidence: BLASTN, 99%,
match=2136)
putative"
/codon_start=1
/protein_id="BAC40816.1"
/db_xref="GI:26354376"
/translation="MAPDIDSCFFLVALFLASCRATTSDDQKICKEVKNKYNCENLGL
NEIPGLTPNTECLFSEFNVLPTQNTFFRLNLITFDLRCQIYWIHEDTFQSOHR
LDLVLVTANPLIFMAETLSPKALHFLFIQTGISIDFPLNKLNSLGLGSNH
ISSIKLPKGPTEKLVLDQFNAAHYLSKEDSSLOQATNLNLNGNDIAGIELGA
EDSAVQSLNFGGTQNLVIFKGLKNSTIQSLWLTGTFEDMDDEIDSPAVFGLCEMSV
FESNLQKHYFNITSSNTHCFGLQELDLTSLHSELPSGLVGLTKLVLANKFE
NLQCIKASNPFSPLTHLSIKNTKRLGLEGTGLENLENRLDLSDHDDTDCNQLQ
RNLSHLSQSLNSLYNEPLSKTEAPKCEPQLLELLDAPTRLVKQKQSPFNHLHLKVL
NLSHSLDIDSSEQLFDGLPALQHLNQGHPKGNIOKTNLSQTLGLRLEILVLSFCDL

FEATURES /
source

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
Trust/MRC building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/
Location/Qualifiers
1. 2157
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM DB:F630107B15"
/db_xref="taxon:10090"
/clone="F630107B15"
/cell_type="NOD-derived CD11c +ve dendritic cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
134. 2119
/note="unnamed protein product; lymphocyte antigen 78
(MGD|MGRI:1194924, GB|NM_008533, evidence: BLASTN, 99%,
match=2136)
putative"
/codon_start=1
/protein_id="BAC40816.1"
/db_xref="GI:26354376"
/translation="MAPDIDSCFFLVALFLASCRATTSDDQKICKEVKNKYNCENLGL
NEIPGLTPNTECLFSEFNVLPTQNTFFRLNLITFDLRCQIYWIHEDTFQSOHR
LDLVLVTANPLIFMAETLSPKALHFLFIQTGISIDFPLNKLNSLGLGSNH
ISSIKLPKGPTEKLVLDQFNAAHYLSKEDSSLOQATNLNLNGNDIAGIELGA
EDSAVQSLNFGGTQNLVIFKGLKNSTIQSLWLTGTFEDMDDEIDSPAVFGLCEMSV
FESNLQKHYFNITSSNTHCFGLQELDLTSLHSELPSGLVGLTKLVLANKFE
NLQCIKASNPFSPLTHLSIKNTKRLGLEGTGLENLENRLDLSDHDDTDCNQLQ
RNLSHLSQSLNSLYNEPLSKTEAPKCEPQLLELLDAPTRLVKQKQSPFNHLHLKVL
NLSHSLDIDSSEQLFDGLPALQHLNQGHPKGNIOKTNLSQTLGLRLEILVLSFCDL

FEATURES /
source

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
Trust/MRC building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/
Location/Qualifiers
1. 2157
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM DB:F630107B15"
/db_xref="taxon:10090"
/clone="F630107B15"
/cell_type="NOD-derived CD11c +ve dendritic cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
134. 2119
/note="unnamed protein product; lymphocyte antigen 78
(MGD|MGRI:1194924, GB|NM_008533, evidence: BLASTN, 99%,
match=2136)
putative"
/codon_start=1
/protein_id="BAC40816.1"
/db_xref="GI:26354376"
/translation="MAPDIDSCFFLVALFLASCRATTSDDQKICKEVKNKYNCENLGL
NEIPGLTPNTECLFSEFNVLPTQNTFFRLNLITFDLRCQIYWIHEDTFQSOHR
LDLVLVTANPLIFMAETLSPKALHFLFIQTGISIDFPLNKLNSLGLGSNH
ISSIKLPKGPTEKLVLDQFNAAHYLSKEDSSLOQATNLNLNGNDIAGIELGA
EDSAVQSLNFGGTQNLVIFKGLKNSTIQSLWLTGTFEDMDDEIDSPAVFGLCEMSV
FESNLQKHYFNITSSNTHCFGLQELDLTSLHSELPSGLVGLTKLVLANKFE
NLQCIKASNPFSPLTHLSIKNTKRLGLEGTGLENLENRLDLSDHDDTDCNQLQ
RNLSHLSQSLNSLYNEPLSKTEAPKCEPQLLELLDAPTRLVKQKQSPFNHLHLKVL
NLSHSLDIDSSEQLFDGLPALQHLNQGHPKGNIOKTNLSQTLGLRLEILVLSFCDL

FEATURES /
source

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
Trust/MRC building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/
Location/Qualifiers
1. 2157
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM DB:F630107B15"
/db_xref="taxon:10090"
/clone="F630107B15"
/cell_type="NOD-derived CD11c +ve dendritic cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
134. 2119
/note="unnamed protein product; lymphocyte antigen 78
(MGD|MGRI:1194924, GB|NM_008533, evidence: BLASTN, 99%,
match=2136)
putative"
/codon_start=1
/protein_id="BAC40816.1"
/db_xref="GI:26354376"
/translation="MAPDIDSCFFLVALFLASCRATTSDDQKICKEVKNKYNCENLGL
NEIPGLTPNTECLFSEFNVLPTQNTFFRLNLITFDLRCQIYWIHEDTFQSOHR
LDLVLVTANPLIFMAETLSPKALHFLFIQTGISIDFPLNKLNSLGLGSNH
ISSIKLPKGPTEKLVLDQFNAAHYLSKEDSSLOQATNLNLNGNDIAGIELGA
EDSAVQSLNFGGTQNLVIFKGLKNSTIQSLWLTGTFEDMDDEIDSPAVFGLCEMSV
FESNLQKHYFNITSSNTHCFGLQELDLTSLHSELPSGLVGLTKLVLANKFE
NLQCIKASNPFSPLTHLSIKNTKRLGLEGTGLENLENRLDLSDHDDTDCNQLQ
RNLSHLSQSLNSLYNEPLSKTEAPKCEPQLLELLDAPTRLVKQKQSPFNHLHLKVL
NLSHSLDIDSSEQLFDGLPALQHLNQGHPKGNIOKTNLSQTLGLRLEILVLSFCDL

FEATURES /
source

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
Trust/MRC building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/
Location/Qualifiers
1. 2157
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM DB:F630107B15"
/db_xref="taxon:10090"
/clone="F630107B15"
/cell_type="NOD-derived CD11c +ve dendritic cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
134. 2119
/note="unnamed protein product; lymphocyte antigen 78
(MGD|MGRI:1194924, GB|NM_008533, evidence: BLASTN, 99%,
match=2136)
putative"
/codon_start=1
/protein_id="BAC40816.1"
/db_xref="GI:26354376"
/translation="MAPDIDSCFFLVALFLASCRATTSDDQKICKEVKNKYNCENLGL
NEIPGLTPNTECLFSEFNVLPTQNTFFRLNLITFDLRCQIYWIHEDTFQSOHR
LDLVLVTANPLIFMAETLSPKALHFLFIQTGISIDFPLNKLNSLGLGSNH
ISSIKLPKGPTEKLVLDQFNAAHYLSKEDSSLOQATNLNLNGNDIAGIELGA
EDSAVQSLNFGGTQNLVIFKGLKNSTIQSLWLTGTFEDMDDEIDSPAVFGLCEMSV
FESNLQKHYFNITSSNTHCFGLQELDLTSLHSELPSGLVGLTKLVLANKFE
NLQCIKASNPFSPLTHLSIKNTKRLGLEGTGLENLENRLDLSDHDDTDCNQLQ
RNLSHLSQSLNSLYNEPLSKTEAPKCEPQLLELLDAPTRLVKQKQSPFNHLHLKVL
NLSHSLDIDSSEQLFDGLPALQHLNQGHPKGNIOKTNLSQTLGLRLEILVLSFCDL

SSIDQHAFTSLKMMHVDLSHNRILTSSIEALSHLKGIVLYNLASNRISIIPLSLPIL
 SQORTNLRONPLDCTCSNIYFLEWYKENMKLEDTEDTLCENPPLLRGLRVLSDVTL
 CSMAAGVIFPLIVFLVFAILLIFAVKYFLRWKYQHI"

ORIGIN

Alignment Scores:

Pred. No.: 0.188 Length: 2157
 Score: 117.00 Matches: 114
 Percent Similarity: 31.72% Conservative: 76
 Best Local Similarity: 19.03% Mismatches: 190
 Query Match: 3.02% Indels: 221
 DB: 4 Gaps: 22

US-10-783-417-2 (1-735) x AK089255 (1-2157)

QY 201 AsnValHisAsnAspPheIleArgGluIleProGly----- 212
 DB 245 AACTGTGAATAATTAGTCTCAATGAATAATCTCGGCACTTTACCAAACTCAACAGAAATGT 304
 QY 213 -----PheGlnLeuGluThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAa 230
 DB 305 TTGAGGTTGAGCTTTAATGCTTCCTACCATCAAAACAGCCTTCAGAGACTTATC 364
 QY 231 AsnPheHisLeuAsnLeuGlnGlnGlyAlaGluLeuAlaAspGlu----- 246
 DB 365 AATCT-CACCTTCTCGATTAAACAGGTGCCAGATTACTGGATACATGAAGATATT 423
 QY 246 ----- 246
 DB 424 CCAAGCCCAACATCGGTTAGACACACTTGTACTAAACCGCAAAATCCCTGATATTATGGC 483
 QY 247 -----TTPAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGly 261
 DB 484 AGAGACAGCACTTAGTGGGCTAAGCAATTGAAGCATCTCTTCATCCAAACAGGAAT 543
 QY 262 -ThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCy 281
 DB 544 ATCCAGTATTGATTATCCCACTGCACAAATCAAAACCTTGAA---AGTCTCTATCT 600
 QY 281 salaaenThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpse 301
 DB 601 CGAAGCAACCAATATTTCCTCCATTAAGCTCCCAAGGTTTCCCAACAGAGAAGCTGAA 660
 QY 301 rIlePhe-----AsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrI 318
 DB 661 GTCTCTGGATTTTCAGATAAATCTATCCATTAACCTGCTCTAAAGAGATATGAGCTCT 720
 QY 318 eSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluVa 338
 DB 721 CCAGCAGCCCACTATCTGAGCCCTTAACCTTAATGAATGACATTCGAGGAATAGAGCT 780
 QY 338 l----- 338
 DB 781 GGGGGCTTCGACTCAGCTGTCTTCCAAAGTTTAAACTTTGGAGGACTCAGAACTTGCT 840
 QY 339 -----LysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThrGluLeAs 355
 DB 841 AGTTATCTTCAAGGGTTTGAAGAACTCTACGATCCAGTCTCTTTGGCTGGGACATTTGA 900
 QY 355 n-----PheAspArgLeuProGlnLeuArgVa 364
 DB 901 GGACATGGATGACGAAGATATTAGTCTCGCGGTGTGGGGTCTCTGTGAATAATGTCTGT 960
 QY 364 lGlnProAsnLeuAlaThrMetGluTyr-----AsnLeuThrArgAlaSerPheLysLe 382
 DB 961 GGAGAGCATCAACCTACAGAGCATATTATCTTCAACATTTCTCCAAACATATCCATTC 1020
 QY 382 uPheSerPheLeuGluGlnPhe---IlePheTyrThrGluAsnThrAsnPheGlyAsnAr 401
 DB 1021 CTTAGTGGCTCCAGAACTGGACCTAAACAGCACTCACCTGAGTGAATGGCTTCTGG 1080
 QY 401 gLeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLe 421

Db 1081 ACTTGTGGGACTAAGC----- 1096
 QY 421 uTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGlu-SerTyrL 441
 Db 1097 -----ACACTTAAGAAATAGTTCTCAGTGC 1122
 QY 441 ysValSerIle---ValThrAspArgGlnSerProProValSerProIleGlnProHisP 460
 Db 1123 AATAAGTTTGAGAAATTTGTCCTCAATCAGTGTCTCCAAATTTCCCTCC-CTTACTCAC 1181
 QY 460 heIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsnThrLeuLysTyrS 480
 Db 1182 TTTCCATC-----AAGGCAACAGAGAGACTTG 1211
 QY 480 erAlaGly---GlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPhePro---- 497
 Db 1212 AGTCGGTACTGGCTGTTTAGAAAACCTAGAAAATCTCCGGAACTTGACCTCAGCCATG 1271
 QY 498 -----ArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheA 515
 Db 1272 ATGACATTGAACTTCTGATTGT-----TGCACCTGCAACTCC 1310
 QY 515 snAsnTyrSerHisIleLeuSer-----HisPheSerLeuPheT 528
 Db 1311 GAAACCTGTCTCACTTACAGAGCCTGAACCTTGAGCTACAATGAACCCCTGAGCCTCAAGA 1370
 QY 528 hrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGly----- 543
 Db 1371 CGGAGGCATTCAAAGAATGCCCTCAGCTAGAACTCTAGATTGGCAATTACTCGACTAA 1430
 QY 544 -----ValLeuGlyTpt 548
 Db 1431 AGGTGAAGATGCACAGAGTCCCTTCCAGAACCTCCATCTTTTGAAGGTGCTGAATCTCT 1490
 QY 548 hrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProA 568
 Db 1491 CCACAGCCTCCTTGAC-----ATCAGCAGTGAAGCAGCTCTTCGATGCCCTGCCAG 1541
 QY 568 laIle-----LysGlyAsnAsnLeuAspThrA 577
 Db 1542 CACTCCAGCATTTGAACCTTACAGGGAATCACTTTCCAAAAGGGAATATTCAAAAGACCA 1601
 QY 577 snSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerG 597
 Db 1602 ACTCA-----CTTCAGACAC 1616
 QY 597 lngLysArgLeuGluIleThr-----CysGluThrProAsnSerThrGlnSerT 613
 Db 1617 TGGGAAGACATAGAAATCCTGTTTATCTCTTTGTGATCTCTCTCCATTCACGACGACG 1676
 QY 613 yrPheIleArgLeuArgTyrAlaThr----- 621
 Db 1677 CTTTCCAGCTCTGAAGATGATGATCATGTAGACTGAGTCACACAGCGGTGACATCCA 1736
 QY 622 -----AsnGlyAlaGlyAsnThrL 628
 Db 1737 GTAGCATTCAGGCTCTTAGTCATCTTAAGGGGATCTACCTCAATCTGGCTCCCAAT--- 1792
 QY 628 euProAsnIleSerLeuThrIleProGlyValIleGlyIleProGlnArg 645
 Db 1793 -----CGCATCAGCATCATCTCTACCCAGTCTCTCTCCCATCTTGTGCCAGCAG 1840

RESULT 9

LOCUS AJ538005 1178 bp mRNA linear EST 01-JUL-2004
 DEFINITION AJ538005 1A1 Timarcha balearica cDNA clone Timarcha6B7, mRNA
 sequence.
 ACCESSION AJ538005
 VERSION AJ538005.1 GI:40311007
 KEYWORDS EST.
 SOURCE Timarcha balearica
 ORGANISM Timarcha balearica
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 Chrysomeloidea; Chrysomelidae; Chrysomelinae; Timarchini; Timarcha.
 1 (bases 1 to 1178)
 Gomez-Zurita, J., Koplaku, F., Theodorides, K. and Vogler, A.P.
 Resources for a phylogenomic approach in leaf beetle (Coleoptera)
 systematics
 (in) Jolivet, P., Santiago-Blay, J.A. and Schmitt, M. (Eds.); NEW
 DEVELOPMENTS IN THE BIOLOGY OF CHRYSOMELIDAE, 19-35 (2004)
 Contact: Gomez-Zurita J
 Department of Entomology
 The Natural History Museum
 Cromwell Road, London SW7 5BD, United Kingdom.
 Location/Qualifiers
 1..1178
 /organism="Timarcha balearica"
 /mol_type="mRNA"
 /db_xref="taxon:79517"
 /clone="Timarcha687"
 /dev_stage="adult"
 /clone_lib="1A1"
 /note="country: Spain;Balearic Islands, Mallorca,
 Esporles"

ORIGIN

Alignment Scores:

Pred. No.: 0.12 Length: 1178
 Score: 115.00 Matches: 75
 Percent Similarity: 34.92% Conservatives: 35
 Best Local Similarity: 23.81% Mismatches: 106
 Query Match: 2.97% Indels: 99
 DB: 1 Gaps: 15

US-10-783-417-2 (1-735) x AJ538005 (1-1178)

QY 470 AsnGlySerSerAsnAsnThrLeuLysTyrSerAlaGlyGly----- 483
 DB 227 AATGCTCAGCGAAACAC-----AATTATCTCTTAATAATAATATCAAGTAATGCT 280
 QY 484 -----SerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPhePro---Arg 498
 DB 281 AATTATTACCAAGTAACGGGAATCATCAAGGAACAATGGGACAGATTCCATCATCG 340
 QY 499 LysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnTyrSer 518
 DB 341 AATAACGATAGATATCAAGAGTACACGGAATCCCTTTCCAAACATGAATAATACGGG 400
 QY 519 HisLLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGln 538
 DB 400 ----- 400
 QY 539 lLeLeuAspThrGlyValLeuGly-----TrpThrHisSerSerValAspArgTyrAsn 556
 DB 401 -----GACCAGGGGAATATGGAATGCTATCCCATTCAGGAATAGTGATACAC 454
 QY 557 AlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnLeuAspThr 576
 DB 455 AATAATAATCCCAAGA-----AATAATCTGCGAGGAATAATTGGAATTT 499
 QY 577 AsnSerLysValIleGlyProGlyHisThrGlyAsnLeuValTyrLeuGlnSer 596
 DB 500 AATTCTGCACCAAGTAATGATATTCAGGAATAGTGGTCATGCTTATCCATCACAAGA 559
 QY 597 Gln-----GlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyr 613
 DB 560 AAAAATTATCCAGGAATAATTTGGAATAGTCTCTTCACCAAGAAAT---AATAGCTAT 616
 QY 614 PheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeu 633
 DB 617 -----ACAGGAATAGTGGCAATGCTTATCCATCACAACAAAT 655
 QY 634 ThrIleProGly-ValIleGlyIle---ProProGlnArgLeuAsnAsnThrPheSerGly 652
 DB 656 AATTATCCAGGAATAATTTGGAATACTTCCACCTTCACGGAATAATACTTATCCAGGA 715

QY 653 ThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPhe----- 666
 DB 716 ACTAATGGGAATACTTACCAAAATCGGAATGGTGAATACTATCCAAAGAATAATGGTAAT 775
 QY 667 GlnPheProSerThrValThr----- 673
 DB 776 AATTTCTCTTCAGTGAATACCAAGTTCATATCCAGGAAGCAGTGCGCAATAATTATCCAGGC 835
 QY 674 LeuProLeuAsnArgAsnIleProPheIlePheAsn-----ArgAlaAspValSer 690
 DB 836 TTCAATATAATAATAATAATATTATTCAGTGCATAACGCCAAGCGGAATACCTGATTATGCT 895
 QY 691 AsnSerIleLeuIleLeuAspLysIle-----GluPheIleProIleThr 705
 DB 896 GGCAGTGTGGGAATACGACAGCAAGTTCGGGAAGTTCGGAAGTCAAAATCCCTAGCACT 955
 QY 706 -----SerSerMetHisGlnAsn----- 711
 DB 956 CAACCAAAAATAATTTACTGGCATAGCGGTAGTGGGCGCACTTACATAAAAATAATATGGG 1015
 QY 712 -----ArgGluLysGlnLysLeuGluThrIleGlnThr 722
 DB 1016 ACGCCCAATGAAGAAGAGTGTGAGTAATGTTAAACG 1054

RESULT 10

CL960568

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-783-417-2 (1-735) x CL960568 (1-2322)

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-783-417-2 (1-735) x CL960568 (1-2322)

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-783-417-2 (1-735) x CL960568 (1-2322)

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-783-417-2 (1-735) x CL960568 (1-2322)

Qy	256	IleGluProAsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIlePro	275
Db	628	TTAGAATTTTGCAGCAGGAACAAAAGAGAACTGGGCTGAGTTCGGAACGACAGTAGACAGT	687
Qy	276	LysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuAspAspGlu	295
Db	688	AGGATCTTAAATCTGCTCAAGGATAGTTTGGCAACAGTATTTGGATGCAATTCGTGATGAC	747
Qy	296	ProAsnMetLysTyrTrpSerIle---PheAsnAspTyrArgArgTyrMetThrIleThrVal	314
Db	748	AACTTGCAGAGTATTACCGTGTTCGAATTTACAGAAATATTCTAAGATCATGCATCTCGTTGCA	807
Qy	315	LeuAspThrIleSerGlnPheSerLeuTyrAspIle-----	326
Db	808	CTTCTCACTGATGAAGATTTTGTTTTTCAGAAAATAGTGTCTTCTGTGTGTCCAAAATA	867
Qy	326	-----	326
Db	868	ACTACTGAATCTTGTCATATCATGTCGTCCCCCTACTAATGTGGGTGCCTCATTCGAAT	927
Qy	327	LysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIle-----LysAsnGlu	344
Db	928	AAAAGTTTAGAAGTTCATGGGCCCAAGTTGATTACCGAAATATCATCTTTACAAATGAG	987
Qy	345	LeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeuArgVal	364
Db	988	CTGGAGGAGGCTGTTAAAGAGCTCATCATTTGGAAATGTGCACATGTGCGATCTTCAAAA	1047
Qy	365	GlnProAsn-----LeuAlaThrMetGluTyr	373
Db	1048	TGTCCAANTGCACATTTTAGCACTGCGAATTCGACCTGGTTGAAGTTGGCATTTCCAGAAATG	1107
Qy	374	AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr	393
Db	1108	AGGATATTACAGAGCTTCCTTTTGTTCACATTTTCAGTTTCGAAGACTTGTATATCTGCTA	1167
Qy	394	GluAsnThrAsnPheGlyAsnArgLeu---ValGlyIleSerAsnArgAspAlaProThr	412
Db	1168	CTGACTTGTCCATGGATTAATGAATTTGATTTGTCTATTATACGACCATTAGACAA	1227
Qy	413	TyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySerProThrThrLys	432
Db	1228	ATGCATTCAATTATCTCTCCAGATTTGAAGGACGG-----GGCGCAGTGAAGCCAAAG	1281
Qy	433	ThrIleArgProPheGluSerTyrLysValSerIleValThrAspArgGlnSerProPro	452
Db	1282	CTCACAGAATATTAT-----GCACAGGATCCGCTA	1311
Qy	453	ValSerProIleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySer	472
Db	1312	TGTGATACCACAATGAATCATATATCTCAAAATATATCAAAATTGATTTGGAAGGCAGA	1371
Qy	473	SerAsnAsnThr-----LeuLysTyrSerAlaGlyGlySerLeuSerAsnTyr	498
Db	1372	AATGATATTACTGATGGACCTGTTCAAGATATCCATTCGAAGAATCTCTTTTGTAT	1431
Qy	489	GlnAsnThrThrPhePheGlnPheProArgLysLysAspCysAsnLeuValIleAspPro	508
Db	1432	ATAACATT-----AAAAATTGCACCTATTTCACAGATGAT	1467
Qy	509	GlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThr	528
Db	1468	GGTATT-----GGTTATGCGGCTATGCTCTCAGCTAATGATGTAATGAATATA-----	1515
Qy	529	TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThr	548
Db	1516	ACAAACTTCCTATGTTTAAAGGGAGACAACACTTACCGATCGCGCTCTCTGCAACTTTGTT	1575
Qy	549	HisSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIle-----	566
Db	1576	GGCTCTTCACCTGAG---TTCCTTGATATTTCGAGACCGGTGGTTTCCATGGTATCATTTA	1632
Qy	567	---ProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGluGlyProGly	585

[illegible]

gene 1..2460
/gene="At5g20250"

ORIGIN

Alignment Scores:
Pred. No.: 0.397 Length: 2460
Score: 115.00 Matches: 170
Percent Similarity: 31.70% Conservative: 109
Best Local Similarity: 19.32% Mismatches: 269
Query Match: 2.97% Indels: 332
DB: 4 Gaps: 49

US-10-783-417-2 (1-735) x CNS09YP2 (1-2460)

```

Qy 57 ThrGlnTyrGlyAspAsnPheGlu-----ThrPheAlaSerAlaAspThrIle 72
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 116 ACGAGGTACAGAAACGAAGAGAGAGAGATGATGAGATTAAACCGGCG-----166

Qy 73 AlaAlaValSerAlaGlyThrIleVal-----SerGlyThrLeuAlaGlyIleGly 90
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 167 GTTCGTATTTCGACCGCGAATCTAATCATCAAAAACCGGAGATTCTAACCGGTGTACCA 226

Qy 91 Gly-----LeuThrSerIleSer-----GlyProIle-----IleGly 103
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 227 GATAATGTCATCAGACGTCAGCATCGGAAGCTCGACCGGTAGAGAGATCTTCGCGGA 286

Qy 104 AlaIle-----IleIleSerPheGlyThrLeuIleThrVal 115
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 287 GCTGTATTTAACAGGAAGAGAGTAACACATCTGACCGATCGGTACGCTTCGCATTC 346

Qy 116 PheTrpProAlaGlyGluInAspLysThrValTrpThrGlnPheIleLysMetGlyGlu 135
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 347 CGGTTTATGTCCTTTCGCGTTTAAG--CTCTGGTGGATGGGCCAGAGAATGGGAGAA 403

Qy 136 IlePheValAspThrPro-----LeuThrGluSer-----145
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 404 ATGGGCCGAGATATTCCCTACGAGACGCAAGTCTTATTGGTTCGAGAGCAACGATGGGTCC 463

Qy 146 -----1leLysGlnLeuLys 150
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 464 CACCTTGAGTCTGACGGAGCTAACCGGCGTCGAGGTAAACGAAAGTTTACACGGTTTTC 523

Qy 151 LeuGlnThrLeuGluGly---PheArgGlnIleLeuGln-----162
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 524 TTGCGGTTAATCGAAGGATCTTCTTCGTTCTGTCGTCTCCAAGAAACGTTAACGATGAGTT 583

Qy 163 -----SerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLys 175
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 584 GAGCTCTGTTTGGAGAGTGGTGCAGTGGACACTAAACCGCTGCTGTTTACTCACTCTCTG 643

Qy 176 ArgLeuGlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeu 195
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 644 TATATTCAACGCC---GGTACAGATCCCTTTTCAGACAATAACGGACGCTATTTCGCACCGTT 700

Qy 196 LysIleArgPheGluAsnValHisAsnAspPheIleArgGluIleProGly-----212
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 701 AAGTTGCATCTGATAGTTCCTGTCACGTCACGAAAGAAAGCTTCCAGGGATGCTTGAC 760

Qy 213 ---PheGlnLeuGluThrTyrLysThrLeuLeuLeuProLetyrAlaGlnAlaAsn 231
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 761 TACTTCCGATGGTGCACTTGGGACCGCTTTTATCAAGAAGTGACTCAGGAAGGGCGTCGAA 820

Qy 232 PheHisLeuAsnLeuLeuGlnGlnIleValaGlu-----LeuAlaAsp 245
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 821 GCTGGTCTTAAGTCTCTCGCGCGCGGTGGTACGCCCGCGAAGTTTGTATCATAGACGAC 880

Qy 246 GluTrpAsnAlaAspIleHisProSerGlnIleGlu-----257
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 881 GGTGGCAATCAGTTTGAGAGAGATGCCACCGTGGAGGCCGAGATGAGAAGAAGTCA 940

Qy 258 -----Pro 258
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 941 CCGATTTTTCGACTGACGGGATCAAGGAGAAACGAAAGTTTAAAGAAGAAGGATGATCCA 1000

```

QY 531 TyrValIleGlyLeuGlnLeuGlnIleLeuAaspThrGlyValIleGly----- 546
Db 1829 -----ACTGAGTTCCTGGCGTGATATAACTGC 1855
QY 547 -----TrrThrHisSerSerValAaspArgTyrAsnAlaIle---SerAspLys 561
Db 1856 CAAGGAGCAGCTTGG-----AGCAGCACAGAGAGAAAAAATTTCCACGAGACTAAA 1909
QY 562 IleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAaspThrAsnSerLysValIle 581
Db 1910 ACTGATAGCCTCACTGGCTCATTCGGTGGTGATGTGATTCATTAATATCGGAGCGCTCC 1969
QY 582 GluGlyProGlyHisThrGlyGlyAsn---LeuValTyrLeuGlnSerGlnGlyArgLeu 600
Db 1970 ACTGATCCACACACCTGATGAGATGCTGCTTTACTCCAGACAGAGCGCACTT 2029
QY 601 GluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAla 620
Db 2030 ATTGTTATGCCATACATATGTCTCTCCAGTCTCACTCAAAATCCGTGAGCAGATC 2089
QY 621 ThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGly 640
Db 2090 TTC-----ACGGTGAGCGCCCATTAGTCATCTTGTGTGATGTGATCTTTT 2134
QY 641 IleProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyr 660
Db 2135 GCCCCAAATGGCTAGTAAACATGTACAATTCGGAGGAGCTATCGAAGACTTAGATAT 2194
QY 661 -----GlyAspPheGlyTyrPhe 666
Db 2195 GAAGCCGAGAGATGAAGTGGTAATGAAGTAAAGATGTGCAAAATTCGGATCTTAC 2254
QY 667 GlnPhe-----ProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePhe 684
Db 2255 TCTTCTGTGAAGCTAAGAGATCGTGTTCAGTCAAAAT---GAGATTGCATTCGAGTAC 2311
QY 685 AsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIleGluPheIleProIle 704
Db 2312 GAT-----TCTCTCTGGATTTGGTCACCTTTGATAGACAAATTCCTTATA 2359
QY 705 ThrSerSerMetHisGlnAsnArgGlyLysLeuGlnIleGlnThrIleLysIle 724
Db 2360 -----GAGAACAAACGATTTTCATCTGTATCCAAAGTTGAGTTA 2395

RESULT 12
CF951415
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CF951415 820 bp mRNA linear EST 20-NOV-2003
UI-M-HLO-cnb-o-16-0-UI-ri NIH BMAP_HLO Mus musculus cDNA clone
IMAGE:30633855 5', mRNA sequence.
CF951415.1 GI:38467284
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers

source
1. 820
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30633855"
/tissue_type="Upper Head"
/dev_stage="9.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP_HLO"
/note="Organ: Head; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CGAATGAAT. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH)."

ORIGIN
Alignment Scores:
Pred. No.: 0.0764 Length: 820
Score: 114.50 Matches: 54
Percent Similarity: 44.26% Conservative: 50
Best Local Similarity: 22.98% Mismatches: 106
Query Match: 2.96% Indels: 25
DB: 7 Gaps: 10

US-10-783-417-2 (1-735) x CF951415 (1-820)

QY 188 LeuGlnGlnAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPheIle 207
Db 51 CTCACCTCTCCCTCGCTGATGTGCGCCAGCAGTATGAAGCGTGGCTGCCAAGAACCTC 110
QY 208 ArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeuLeuProIleTyrAla 227
Db 111 CAGAGGCCGAG-----GAATGGTACAAAGTCCAAAGTTGCTGACCTC---TCT 155
QY 228 GlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTrp 247
Db 156 GAGGCTCCCAACCGAACAACGATGCCCTGGCCAGGCCAAGCAGAGTCAACACGATAC 215
QY 248 AsnAlaAspIleHisProSerGlnIleGluProAsnAla-----GlyThrSerAsp 265
Db 216 CGGAGACAGGTGCAGTCACTCACCTGTGAGTGGATGCCCTTAAAGGCACCTAACGAGTCC 275
QY 266 TyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyr 285
Db 276 CTGAGCGCCAGATGCGTGGATGGAAGAAATTTGCCCTTGAAGCTGCTAAC---TAC 332
QY 286 ArgThrGlyLeuLysAsnLeuArgAspGlu---ProAsnMetLysTrpSerIlePheAsn 304
Db 333 CAGGACACTATTGGCCGCTGCAGATGAGATCCAAACATGAGAGAGAGATGCTCGT 392
QY 305 AspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyr 324
Db 393 CACCTTCGTGAATACCAAGATCTCTCAATGTTAAGATGGAAGAGAAATTTTGCCTT--- 449
QY 325 AspIleLysArgTyrArgAspSerIleGlyIle-----GluValLysGlyIleLys 342
Db 450 GAAGCTCTAACTACACGAGCAGTATTGGCCGCTGCAGGATGAGATCCANAACATGAAG 509
QY 343 AsnGluThrThrArgGluIle-----TyrThrThrGluIleAsnPheAspArgLeuPro 360
Db 510 GAAGAGATGGCTCTGCTACCTTCGTGGAATCAAGATCTGCTCAATGTTAAGATGCCCTG 569
QY 361 GlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPhe 380


```

Db      226 ACCCCCTTTTGGGAAAAAACAACACTCTTCCCTTCACTA-AACCTACTTCTTT 168
Qy      681 -----ProPheIlePheAsnArgAlaAspValSerAsnSerIleLeuI 695
Db      167 AATAAAACCCGCCACCTTTTGG-----AAAAATACCTTCTCTCC 123
Qy      695 leileAspLysIleGluPheIleProIleThrSerSerMet 708
Db      122 GCCCTCGCCCTTTTATTATTTACACCCATTACTCTCTCCCTC 82

RESULT 14
CL974457      2841 bp DNA linear GSS 21-SEP-2004
LOCUS      OsIFCC025896 Oryza sativa Express Library Oryza sativa (indica
DEFINITION      cultivar-group) genomic, genomic survey sequence.
ACCESSION      CL974457
VERSION      CL974457.1 GI:52403441
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 2841)
AUTHORS      Ma,L., Wang,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
Wong,K.S., Deng,X.W. and Wang,J.
TITLE      An analysis of transcriptional regulation of the rice genome and
JOURNAL      its comparison to Arabidopsis
COMMENT      Unpublished (2004)
Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
FEATURES
source
1..2841
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/notes="Oryza sativa exon trapped genomic sequences"

ORIGIN
Alignment Scores:
Pred. No.:      0.982      Length:      2841
Score:      112.50      Matches:      175
Percent Similarity:      33.06%      Conservative:      111
Best Local Similarity:      20.23%      Mismatches:      313
Query Match:      2.91%      Indels:      267
DB:      10      Gaps:      47

US-10-783-417-2 (1-735) x CL974457 (1-2841)
Qy      24 ArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGln-----ProLeu 41
Db      499 CGCAATGAAGACGTTGCCAAGAAAATTTGCACACATATAACCATATAGGCTAAATCCCTG 558
Qy      42 GlnAsnThrAsnTyrLysGluTTPLeuAsnMetCysGlnGlyAsnThrGlnTyrGlyAsp 61
Db      559 GACTCTACC-----ATGTGTGGGAATATATAAGCAACGCGAGT 597
Qy      62 AsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleVal 81
Db      598 AATTAGACGTAGAGGTGATAAGACCAATTCGACCAAGTC-----GGATGGGTCAAT 651
Qy      82 SerGlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGlyIle 101
Db      652 GCC-----AAGAAATGTGAGGCGATAGCATTAGCCCATGCACTTGTTTC 699

```

```

Qy      102 Ile-----GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAla 119
Db      700 TTGCTATCTGAATGGACCTTTGTGGAATGGAGACACTG----- 738
Qy      120 GlyGluGlnAspLysThrValTTPThrGlnPheIleLysMetGlyGluIlePheValAsp 139
Db      739 -----AGCAATAGTAGATATTGGACT-----GAAGCTTTTCTGTAT 774
Qy      140 ThrProLeuThrGluSerIleLysGlnLysLeuGlnThrLeuGluGlyPheArgGln 159
Db      775 AATTCAGGTGTGCCATCTCTGAAA---CTAACTTACAAGAATATGCCACCATATTGAGA 831
Qy      160 IleLeuGlnSerTyr-----AsnThrAla----- 167
Db      832 TTATGCTTTTGCCTACTGGCCCATCTTCCCTAAAGGTCAATATAGCTAAAGCAAGTCTG 891
Qy      168 LeuAspAspTTPArgLysLeuLysArgLeuAlaProGlyLeuProProSerSerAla 187
Db      892 GTTCACCAATGGATGCTCTTGGTTTATTGAG-----CCATCTAAAACA 936
Qy      188 LeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPheIle 207
Db      937 TTTTCATCCGTACGACTTGGCGAGAG-----TATATC 969
Qy      208 ArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeuProIleTyrAla 227
Db      970 AGACACACTTGTGGG-----ATGCTTCTTCCTCAGCGTTCATATTGCACACTGAACAA 1023
Qy      228 GlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGlyAlaGluLeuAlaAspGluTTP 247
Db      1024 GAAGTTTTCACCATGCGACGATATGTTTCATGATGTTGCGAGATCGGTATGATGAAGAG 1083
Qy      248 AsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTyrTyr 267
Db      1084 TTGGTTTTCTTATATGACACAAAATTTAGCAGTACCACAGAACAAAATTTTGCACATAT 1143
Qy      268 LysLeuLeuLysGluAsnIleProLysTyrSerAsnTyr----- 280
Db      1144 GCTTTGCTT---GAAAATTATAGCAAGTCATCAAAATCTCAACTATATTATTCCTGCCACG 1200
Qy      281 -----CysAlaAsnThrTyrArgThrGly----- 288
Db      1201 TTAAGGCGCAGTACATACTTCGAATTGCAGCAAAATTTGGTGTCCCAAGGTGATGAATTTCA 1260
Qy      288 ----- 288
Db      1261 TTCACAAAGTTCTTGGTGTCTTAGATTAACTGCTGCTCCATAAGGATCTCTCCCAAGC 1320
Qy      289 -----LeuLysAsnLeuArg-----AspGluProAsnMetLysTrpSerIle 302
Db      1321 TCTATTGGCAAAATGAAGCAGCTGAGGTTTCTTATCGCTCCACACATAGGAGATAATGTG 1380
Qy      303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db      1381 TTCCCA-----AAGAGTATTACTCTGTCTCCCAAACTA----- 1413
Qy      323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyIleGluValLysGlyIleLys 342
Db      1414 -----AAATACCTTGGCTCCATCGATCTTTTCGAATTTTCAGCATTTGCAA 1458
Qy      343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362
Db      1459 GGATCAATTAGCAAAACATGCTTGTCTGNTACATCTTGACCTATCTGGTGTCTCTAATATC 1518
Qy      363 ArgVal---GlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPheLys 381
Db      1519 AGAGTTATACAGCCAGAGAGCTCTATGT-----GSCCTCACCAAACTGCAGTTTGTG 1569
Qy      382 LeuPheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsnArg 401
Db      1570 AACTTATCGTGTGCTCCATCTTCTTCAAAATACTGCGAGAAAATATTGCAAGTCTTACTGAA 1629

```

QY	402	Leu	-----ValGlyIleSerAsn-----ArgAspAlaProThrTySerAsn	415
DB	1630	CTGCAATATTAAACCTTATCAAACTGCTCTCTCTACGCCAGTTTACCATTCTCATATTGGG	1689	
QY	416	ThrIleThrGluThrLeuTyGlyGluArgThGly---SerProThrThrIleThrIle	434	
DB	1690	AGCCTCACTGAACATCAATATTGAACTTATCAGGTTGTGAGGCCCTTGTAAATAATGCCA	1749	
QY	435	ArgProPheGluSerTyrlsValSerIle-----	444	
DB	1750	ATGTCGTTTGGGAATCTGAAGAATTGGTACATCTTGACCTGTGAGCTGCTCGGGGTA	1809	
QY	445	-----ValThrAspArgGlnSerProProValSerPro	455	
DB	1810	CAAGATTCAAAACAAGTCTTTGGTGGGCTCACCAAACTGCAGTACTTAAACTTATCAAG	1869	
QY	456	Ile-----GlnProHisPheIle-----	461	
DB	1870	ATATTGGTAGAACACAGTGGGTGAATAATGGGANGTTTATCTCGTAACATATCAGTACT	1929	
QY	462	IleAsnGlnIleGluLeuTyrlsAsnGlySerSerAsnAsnThrLeuIleTyTyr-----	479	
DB	1930	CTTAATGATCTTGAG---TATCTAAACCTTCTCGGAATGACGAGGATGATATCTACCA	1986	
QY	480	SerAlaGlyGlySerLeuSerAsnTyrlsGlnAsnThrThrPhePheGlnPheProArgLys	499	
DB	1987	AGAAGCCTTGGAAACCTCAAAAAGCTTCAA---ACATTGGACCTTTTCATATGTCGGTCT	2043	
QY	500	-----LysAspCysAsnLeuValIleAspPro	508	
DB	2044	CTGAGATCCCTCCACACAGCATAGAATAATTGACAGTCTGGAGTTTCTAATAGTTGTG	2103	
QY	509	GlyCysSerProAsnPheAsnAsnTyTyr-----Ser	518	
DB	2104	GGGTGCTCAGATCAACTTAAGAGTATTTTGAGAAATCTCATTCAGAATATTCAGTCT	2163	
QY	519	HisIleLeuSerHisPheSerLeuPheThrTyrlsSerTyrlsValIle-----Gly	534	
DB	2164	CATTAC-----CACACTTCATTGTTTCACCTATTTCTGTCCATAGAAGAAGCACGAGGA	2217	
QY	535	LeuGlnLeu-----GlnIleLeuAspThrGlyValLeuGlyTrpThrHisSer	551	
DB	2218	ATTGAGTTGTGAGAAACAAATCTTTCAACATTTGACATTTTCATTGGACT---TCAAGG	2274	
QY	552	ValAspArgTyrlsAsnAlaIleSerAspIleIleThrMetIlePro-----	567	
DB	2275	GCTGACCGACTTTTGGAGGACAAGATGTTTGGGAGAACTTAATGCCACCAAGAGTCTG	2334	
QY	568	-----AlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGluGlyPro	584	
DB	2335	TGGNAATTAGCATACAAGGTATGATAGACACACTTCCCACTTGTTTGTAGGATC	2394	
QY	585	GlyHisThrGlyGlyAsnLeuValTyrlsGlnSerGlnGlyArgLeuGluIleThrCys	604	
DB	2395	TCACATCATCTTCCAAGTCTGTTTAAGATTGAGTT-GAAGGATTTGCACAGATG-----	2447	
QY	605	GluThrProAsnSerThr-----	610	
DB	2448	CAGACCTGCCACCACCTGGGACAGTTGCCAAACCTGGAATGAAGTGCATCTCCAACAGAT	2507	
QY	611	---GlnSerTyrls-----PheIleArgLeuArgTyrlsAlaThrAsnGlyAla	624	
DB	2508	GGACAGTCTCACAAAATTGATGCGGTTCTTGGCGTGGGACAA-----	2552	
QY	625	GlyAsnThrLeuProAsnIle---SerLeuThrIleProGlyValIleGlyIleProPro	643	
DB	2553	---AGGGCCATTCGCAAACTGAAAGTCTTCAACCTTGAGGATAT-----	2594	
QY	644	GlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrlsAsnAsnLeuGln-TyrGly-----	661	
DB	2595	-----GAAACAGTTGGGAAGAGTGGAGCACCACACACACAGTAAGAGTGTGTA	2642	
QY	662	AspPheGlyTyrlsPheGlnPheProSerThrValThrLeuProLeuAsn-----	677	

Db	2643	GGATAGCATAGACTTTCATGTGTTCCCTATGCTAGTTACACTGTCTCAATAGAGCTTTGTCCAAT	2702
Qy	678	-----ArgAsnIleProPheIlePheAnAArgAlaAspValSerAsnSerIleLe	694
Db	2703	GCTTAGACTGAACCACTGCCCACCGATGTTTCACAGCTGGCTCATATCAGAGGCGACAA	2762
Qy	694	ulleIleAsp-----LysIle-----GluPheIlePr	703
Db	2763	GGCTGTCGACACGGGTTTAACGGAGGCTCTGATGAAGATTTTCGCTGAAGATTTTCATGCC	2822
Qy	703	olleThrSerSer	707
Db	2823	TACCGACTCTTCA	2835
RESULT 15			
LOCUS	CNS0ALJE	1143 bp	linear
DEFINITION	Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTFB30ZG08 of Flowers and buds of strain col-0 of Arabidopsis thaliana (thale cress).		
ACCESSION	EX829716		
VERSION	EX829716.1	GI:42458466	
KEYWORDS	HTC; GSUT cDNA.		
SOURCE	Arabidopsis thaliana	(thale cress)	
ORGANISM	Arabidopsis thaliana		
REFERENCE	1	(bases 1 to 1143)	
AUTHORS	Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.		
TITLE	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation		
JOURNAL	Unpublished		
REFERENCE	2	(bases 1 to 1143)	
AUTHORS	Genoscope.		
TITLE	Direct Submission		
JOURNAL	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)		
COMMENT	The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_FF/Full length http://www.genoscope.cns.fr/cgi-bin/ggb/ggbraource=Arabidopsis.		
FEATURES	Location/Qualifiers		
source	1. .1143 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /clone="GSLTFB30ZG08" /tissue_type="Flowers and buds" /ecotype="Col-0" /plasmid="pCMVSPORT 6" complement(1..1143) /gene="At5g56640"		
gene			
ORIGIN			
Alignment Scores:			
Pred. No.:	0.335	Length:	1143
Score:	111.00	Matches:	58
Percent Similarity:	37.33%	Conservative:	23

Best Local Similarity: 26.73% Mismatches: 86
Query Match: 2.87% Indels: 50
DB: 4 Gaps: 11

US-10-783-417-2 (1-735) x CNS0A1JE (1-1143)

```
QY 139 AspThrProLeuThrGluSerIleLeuGlnLeuLysLeuGlnThrLeuGluGlyPheArg 158
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 380 GATCCCGACCTCGACGACACACAAATTCACATTTGTTCAATCTGCGAAGCAATCCGT 439
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 159 GlnIleLeuGlnSerTyrAsn-----ThrAlaLeu---AspAspTrpArgLys 173
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 440 AAAGATTATCTTAACGAAATTTGGCTTCATCTACGCTTATCCATGATCTTTGGGAAG 499
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 174 LeuLysArgLeuGlnAlaProGlyLeuProSerSerIleLeuGlnGlnAlaLeu 193
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 500 GTTCTTACCCTACACAGATTGGAGGGCTTCCTCAATGGGCTGTTGTTGGTGACACATTT 559
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 194 ThrLeuLysIleArgPheGlu-----AsnValHisAsnAspPheIleArgGluIlePro 211
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 560 CTTGTTGGGTGTCATTTGATGATCTAAGTACACCAAGTATTTTATGGAACCCCT 619
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 212 GlyPheGlnLeuGluThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAsn 231
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 620 GATTTTAAACAACCTTAAGTACATACC--AAAGCTGGAATTTATTCTGAAGGCTGTGGA 676
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 232 PheHisLeuAsnLeuGlnGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIle 251
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 677 CTTGAAAATGTTCTTATGTCATGGGGA----- 703
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 252 HisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTyr---TyrLysLeuLeu 270
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 704 -----CATGATGACTACATGATCTCGTGGCC 730
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 271 LysGluAsn-----IleProLys----- 276
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 731 AAGGAGAATGGAAGCACCTTACCATCCCGGGACTATTATCATCCGATACCATTCCTTT 790
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 277 TyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro 296
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 791 TATCCATTGTACAAGGCTGGAGCTTAC--ACTCACCTTATGAACGAGGAAGACAAGGAG 847
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 297 AsnMetLysTrp---SerIlePheAsnAspTyrArgArgTyrMetThrIleThrVal--- 314
   ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : |||
Db 848 AATCTCAAAATGGCTTCATGTTTTTAACAAGTATTACTTTATACAGCAAGAGCAAGTACAC 907
   ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : |||

QY 315 -----LeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyr 329
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 908 GTTAATGTCGAGAAGGTTAAGCCGTACTACATTTCTTATCAAGAAATAT 958
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: December 12, 2005, 04:21:08
Job time : 7319 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 11, 2005, 14:08:26 ; Search time 332 Seconds
(without alignments)
3935.265 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869

Sequence: 1 MNQNDNNEVEIIDSHTSPY.....KLETIQKINTFFNHTKTL 735

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10783417/runat_01122005_141954_24196/app_query.fasta_1.903
-DB=Issued Patents NA -OFMT=fastap -SUFIX=tni -MINMATCH=0.1 -LOOPECL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=spt -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10783417 @CNG 1_1_200 @runat_01122005_141954_24196 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1060.5	27.4	3543	3	US-09-224-024-27
2	1060.5	27.4	3543	6	PCT-US94-07902-27
3	899.5	23.2	2061	3	US-09-224-024-30
4	899.5	23.2	2061	6	PCT-US94-07902-30
5	775.5	20.0	3507	2	US-08-315-468-3
6	676	17.5	1953	2	US-08-315-468-5
7	676	17.5	1953	3	US-07-941-650A-3
8	663.5	17.1	2430	6	PCT-US92-00040-1
9	663	17.1	1959	3	US-08-996-441B-37

10	663	17.1	1959	3	US-08-993-722A-37	Sequence 37, Appl
11	663	17.1	1959	3	US-08-993-170A-37	Sequence 37, Appl
12	663	17.1	1959	3	US-08-993-775B-37	Sequence 37, Appl
13	663	17.1	1959	3	US-09-427-770-37	Sequence 37, Appl
14	663	17.1	1959	3	US-09-427-769-37	Sequence 37, Appl
15	660.5	17.1	3414	2	US-07-973-320-3	Sequence 3, Appl1
16	660	17.1	1959	3	US-09-377-466B-3	Sequence 3, Appl1
17	660	17.1	1959	3	US-10-232-665-3	Sequence 3, Appl1
18	658	17.0	1959	3	US-08-996-441B-27	Sequence 27, Appl
19	658	17.0	1959	3	US-08-996-441B-35	Sequence 35, Appl
20	658	17.0	1959	3	US-08-993-722A-27	Sequence 27, Appl
21	658	17.0	1959	3	US-08-993-722A-35	Sequence 35, Appl
22	658	17.0	1959	3	US-08-993-170A-27	Sequence 27, Appl
23	658	17.0	1959	3	US-08-993-170A-35	Sequence 35, Appl
24	658	17.0	1959	3	US-08-993-775B-27	Sequence 27, Appl
25	658	17.0	1959	3	US-08-993-775B-35	Sequence 35, Appl
26	658	17.0	1959	3	US-09-427-770-27	Sequence 27, Appl
27	658	17.0	1959	3	US-09-427-770-35	Sequence 35, Appl
28	658	17.0	1959	3	US-09-427-769-27	Sequence 27, Appl
29	658	17.0	1959	3	US-09-427-769-35	Sequence 35, Appl
30	656	17.0	1962	3	US-09-377-466B-5	Sequence 5, Appl1
31	656	17.0	1962	3	US-10-232-665-5	Sequence 5, Appl1
32	656	17.0	2050	3	US-08-996-441B-101	Sequence 101, App
33	656	17.0	2050	3	US-08-993-722A-101	Sequence 101, App
34	656	17.0	2050	3	US-08-993-170A-101	Sequence 101, App
35	656	17.0	2050	3	US-08-993-775B-101	Sequence 101, App
36	656	17.0	2050	3	US-09-427-770-101	Sequence 101, App
37	656	17.0	2050	3	US-09-427-769-101	Sequence 101, App
38	655	16.9	1959	3	US-08-996-441B-17	Sequence 17, Appl
39	655	16.9	1959	3	US-08-996-441B-59	Sequence 59, Appl
40	655	16.9	1959	3	US-08-993-722A-17	Sequence 17, Appl
41	655	16.9	1959	3	US-08-993-722A-59	Sequence 59, Appl
42	655	16.9	1959	3	US-08-993-170A-17	Sequence 17, Appl
43	655	16.9	1959	3	US-08-993-170A-59	Sequence 59, Appl
44	655	16.9	1959	3	US-08-993-775B-17	Sequence 17, Appl
45	655	16.9	1959	3	US-08-993-775B-59	Sequence 59, Appl

ALIGNMENTS

RESULT 1

US-09-224-024-27
; Sequence 27, Application US/09224024
; Patent No. 6056953
; GENERAL INFORMATION:
; APPLICANT: Leslie Hickie
; APPLICANT: Jewel Payne
; TITLE OF INVENTION: Materials and Methods for the Control of
; TITLE OF INVENTION: Calliphoridae Pests
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/224,024
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/856,226
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA79

TELECOMMUNICATION INFORMATION:

TELEPHONE: 904-375-8100

TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 3543 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-09-224-024-27

Alignment Scores:

Pred. No.: 2,54e-102 Length: 3543
 Score: 1060.50 Matches: 271
 Percent Similarity: 50.26% Conservative: 117
 Best Local Similarity: 35.10% Mismatches: 279
 Query Match: 27.41% Indels: 105
 DB: 3 Gaps: 24

US-10-783-417-2 (1-735) x US-09-224-024-27 (1-3543)

QY 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
 DB 1 ATGAATCCTTATCAAAATAAATAAATGAATATGAACATTAAATGCTTCACAAAAAATA 60
 QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
 DB 61 -----AATATATCTAATAATATATCAAGATATCCATAGAAATAGTCCAAACAAATTA 114
 QY 41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
 DB 115 TTACAAAGTACAAATATATAAAGATGGCTCAATATGTGTCAACAGATCAGCAGTATGGT 174
 QY 61 AspAsnGluThrPheAlaSerAlaAspThrIleAlaIleAlaValSerAlaGlyThrIle 80
 DB 175 GGAGATTTTGAACCTTTTATTGTAGT-----GGTGAACCTCAGTGGCTATACTATT 225
 QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
 DB 226 GTAGTTGGGACCGTACTGCTGTTTCGGGTTCACAACA-----CCCTTAGGA 273
 QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGly 120
 DB 274 CTT-----GCTTTAATAGTTTGTGTACATTATACAGTCTCTTTTCAGCCCAA 324
 QY 121 GluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
 DB 325 GACCAATCTAACACA---TGGAGTGACTTTATATAACACAAACTAAAAATATTATAAAAA 381
 QY 141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
 DB 382 GAAATAGCATCAACATATATAAGTAATGCTPAATAAAATTTTAAACAGGTCGTTAATGTT 441
 QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaPro 180
 DB 442 ATCAGCAGCTTATCATATATCACTTAAACAAATGGAG-----AATAATCCA 486
 QY 181 GlyLeuProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGlu 200
 DB 487 AACCCCAAAATACTCAGGATGTAAGGACACAAATCCAGCTAGTTTCATTACCATTTTCAA 546
 QY 201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr 218
 DB 547 AATGTCATTCAGAGCTTGTAAACTCTTGTGCTCCTAATCCTAGTATGCGATTACTAT 606
 QY 219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAsnPheHisLeuAsnLeuGln 238
 DB 607 AACATCATGTATTATCTAGTTATGCAAGCAGCAAACTTACATCTGACTGCTGATTAAAT 666
 QY 239 GlnGlyAlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluPro 258
 DB 667 CAAGCCGTCAAATTTGAGCGTATTTAAAAAACAATCGACAATTCGATTATTATAGACCT 726

QY 259 AsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSer 278
 DB 727 ---TTGCCAACAGCAATTTGATTATATCCAGTATTGACTAAAGCTTATAGAAGATTACACT 783
 QY 279 AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro----- 296
 DB 784 AATTATTGTGTAAACAACCTTATAAAAAAGGATTAATTAATTAATAAACACGCGCTGATAGT 843
 QY 297 -----AsnMetLysTyrSerIlePheAsnAspTyrArgArgTyrMetThrIle 312
 DB 844 AATCTTGATGGAATATATAAAGTGAACACACATACATACGTATCGAACAAATGACTACT 903
 QY 313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer 332
 DB 904 GCTGTATTAGATCTGTGTGCACTCTTTCCTAATTAATGATGATAGGTAAATAT----- 954
 QY 333 IleGlyGlyIleGluValLysGlyIleLysAsnGlnLeuThrArgGluIleTyrThr 352
 DB 955 -----CCAATAGGTGTCCAATCTGAACCTTACTCGAGAAATTTATCAGGTA 999
 QY 353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
 DB 1000 ---CTTAACCTTCGAAGAAAGCCCTCTATAATATTAT-----GACTTCAATATCAAGAG 1050
 QY 373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyr 392
 DB 1051 GATTCACCTACACGTAGACCG---CAATTATTACTTGGCTTGATCTTCTGAATTTTAT 1107
 QY 393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
 DB 1108 GMAAAGCGCAACTTCTCTAATAATTTTTCACCAGCCATTATAATATGTTTCATTAC 1167
 QY 402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
 DB 1168 ACACCTGATAATATATCCAAATATCTAGTGTGTTTGGAAATCAATGTAAGTAAATA 1227
 QY 421 LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440
 DB 1228 TTA-----AAATCTCTGTTGGCAACAATATTATATT----- 1263
 QY 441 LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
 DB 1264 -----TTT 1266
 QY 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
 DB 1267 TTATTAAATGTCATAAGCTTAGATAATAATAATATCTAAATGATTATAATAATTATTAGTAAA 1326
 QY 477 -----LeuLysTyrSerAlaGlyGly 483
 DB 1327 ATGGAATTTTATTAATACTAATGCTAGACTTTTGGGAAGAAGAACTTACAGCAGGATCT 1386
 QY 484 SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysLysAspCys 502
 DB 1387 GGGCAATAACTTATGATGTAATAATAATAATTTTCGGTTACCAATCTTAACCAAGA 1446
 QY 503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
 DB 1447 GAGAATCAAGCAATCCCTACCCCTTTTCCACATATGATACTATAGTCATATTATTATCA 1506
 QY 523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThr 542
 DB 1507 TTTTATTAAGTCTTAGTATCCCTGCAACATATAAAACTCAAGTGTAT----- 1554
 QY 543 GlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
 DB 1555 ---ACGTTTCTGGACACACTCTAGTGTGATCTTAAATAATACAAATTATACACATTTA 1611
 QY 563 IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582
 DB 1612 ACTACCCAAATTCAGCTGTAAAGCGAATTCACCTTGGGACTGCTTCTTAAGTCTTCTCAA 1671

```
QY 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluLeu 602
Db 1672 GGACCTGGTCATACAGAGGGGATTAAT-----GATTCAAAGATCATTTCAAAT 1725
QY 603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
Db 1726 ACATGTCAACACTCAAAATTTTCAACAATCGTATTTTATAGAAATCGTATGCTTCAAT 1785
QY 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db 1786 GGAAGCGCAAAATACAGAGCTGTATTAATCTTTAGTATCCAGGGGTAGCAGAACTG--- 1842
QY 643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnLeuGlnTyrGlyAsp 662
Db 1843 GGTATGGCACTCAACCCACTTTTCTGTCAGATATACGAATTTAAATATATAAGAT 1902
QY 663 PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
Db 1903 TTTCACTACTTAGAATTTCTCAAGAGGTGAATTTGCTCCAAATCAAAACATATCTCTT 1962
QY 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleAspLysIleGluPhe 701
Db 1963 GTGTTTAATCGTTCGGATGTATATACAAACACACACAGTACTTATTTGATAAAATGAAT 2022
QY 702 IleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGln 721
Db 2023 CTCCTCAATTAATCGTCTTATAGAGAGGATAGAGAGAAACAAATTTAGAAACAGTACAA 2082
QY 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db 2083 CAATAATTAATACATTTATGCAATCCTATATAA 2118

RESULT 2
PCT-US94-07902-27
Sequence 27, Application PC/TUS9407902
GENERAL INFORMATION:
APPLICANT: Street address: 4980 Carroll Canyon Road
APPLICANT: City: San Diego
APPLICANT: State/Province: California
APPLICANT: Country: US
APPLICANT: Postal code/Zip: 92121
APPLICANT: Phone number: (619) 453-8030 Fax number: (619) 453-6991
APPLICANT: Telex number:
TITLE OF INVENTION: Materials and Methods for the Control of
TITLE OF INVENTION: Calliphoridae Pests
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/07902
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA79
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 3543 base pairs
```

```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US94-07902-27
Alignment Scores:
Pred. No.: 2,54e-102 Length: 3543
Score: 1060.50 Matches: 271
Percent Similarity: 50.26% Conservative: 117
Best Local Similarity: 35.10% Mismatches: 279
Query Match: 27.41% Indels: 105
DB: 6 Gaps: 24
US-10-783-417-2 (1-735) x PCT-US94-07902-27 (1-3543)
QY 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 1 ATGAATCCTTATCAAAATATAAATGAATATGAACATTAATGCTTCACAAAAAATTA 60
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 61 -----AATATATATCAATAATATACAGATATCAATAGAAAAATAGTCCAAAAACAATTA 114
QY 41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
Db 115 TTAACAAGTACAAATTTATAAAGATTGGCTCAATATGTCTCAACAGAAATCAGAGTATGT 174
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80
Db 175 GGAGATTTTGAACCTTTTATTGATAGT-----GGTGAACCTCAGTCCCTATATATT 225
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGly 100
Db 226 GTAGTGGGACCGTACTGACTGGTTTCGGGTTTCAACA-----CCCTTAGGA 273
QY 101 IleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120
Db 274 CTT-----GCTTTAATAGTTTGGTACATTAATACCACTCTTTTCCAGCCCAA 324
QY 121 GluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db 325 GACCAATCTAACACA---TGGAGTGACTTTTAAACACAAACTATAAATAATATAAAAAA 381
QY 141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
Db 382 GAATAGCATCAACATATATAAGTAATCTAATAAAATTTTAAACAGGTCGTTTAATGTT 441
QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaPro 180
Db 442 ATCAGCACTTATCATATACCTTTAAACATGGGAG-----AATAATCCA 486
QY 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGlu 200
Db 487 AACCCACAAATACTCAGAGTGTAAAGGACACAAATCCAGCTAGTTTCATTACATTTTCAA 546
QY 201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr 218
Db 547 AATGTCATTCAGAGCTGTAAACTCTTGCTCTCTAATCCTAGTATGCGATTAAT 606*
QY 219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGln 238
Db 607 AACATAGTATGATTATCTAGTTATGCAAGCAGCAAACTTACATCTGACTGATTAAAT 666
QY 239 GlnGlyAlaGluLeuAlaAspGluTyrPheAsnAlaAspIleHisProSerGlnIleGluPro 258
Db 667 CAAGCCGTCAAAATTTGAAGCGTATTTTAAAAACAATCGACAAATTCGATTATTAGAGCCT 726
QY 259 AsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSer 278
Db 727 ---TTGCCAACACGCAATTTGATTATATCCAGTATTGACTAAAGCTATAGAAATACACT 783
QY 279 AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro----- 296
```

Db 784 AATTATTGTGTAACAACCTTATAAAGGATTAAATTAATAAAGCAGCGCTGATAGT 843
QY 297 -----AsnMetIysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIle 312
Db 844 AATCTTGATGGAATAATAAATCGAACACATACATACGATCGAACAAATGACTACT 903
QY 313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleIysArgTyrArgAspSer 332
Db 904 GCTGTATTAGATCTTGTCGACTCTTCTCTAATTATGATGATGATGATGATGATGAT 954
QY 333 IleGlyGlyIleGluValIysGlyIleIysAsnGluLeuThrArgGluIleTyrThr 352
Db 955 -----CCAATAGTGTCCTCAATCTGAACTTCTGAGAAATTTATCAGGTA 999
QY 353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
Db 1000 -----CTTAACCTCGAAGAACCCCTATAATATAT-----GACTTTCATATCAAGAG 1050
QY 373 TyrAsnLeuThrArgAlaSerPheIysLeuPheSerPheLeuGluGlnPheIlePheTyr 392
Db 1051 GATTTCATCTACAGTAGACCG-----CATTTATTACTTGGCTTGATCTTTGAAATTTTAT 1107
QY 393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
Db 1108 GAAAGCGCAAACTACTCTAATAATTTTCCACCGCAATTAATATATGTTTCATTAC 1167
QY 402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
Db 1168 ACATCTGATAATATATCCCAAAATCTAGTGTGTTTGGAAATCAATGTAATGATGATAA 1227
QY 421 LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440
Db 1228 TTA-----AAATCTCTGTTGGCAACAAATATTATATAT----- 1263
QY 441 LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
Db 1264 ----- 1266
QY 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
Db 1267 TTATTAAATGTCATAAGCTAGTAAATAAATATCTAAAGATTATATAATATATGATAA 1326
QY 477 -----LeuLysTyrSerAlaGlyGly 483
Db 1327 ATGGATTTTTTATTAATACTAATGCTAGTACTTTTGGAGAAAGAACTTACAGCAGGATCT 1386
QY 484 SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysLysAspCys 502
Db 1387 GGGCAAAATAACTTATGATGTAATAAATAATATTTTCGGGTGTACCAATTTCTTAAACCAAGA 1446
QY 503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
Db 1447 GAGAAATCAAGCAATCCCTACCTCTTTTCCAACTATGATAAATATGATCATATTTTATCA 1506
QY 523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThr 542
Db 1507 TTTATTAAAGCTTAGTATCCCTGCAACATATAAATACTCAAGTGAT----- 1554
QY 543 GlyValLeuGlyTrpThrHisSerSerValAspArgTyrAsnAlaIleSerAspIle 562
Db 1555 ---ACGTTTGCTGGACACACTAGTGTGTGCTTAAATAAATAATTTATACACATTTA 1611
QY 563 IleThrMetIleProAlaIleIysGlyIysAsnAsnLeuAspThrAsnSerLysValIleGlu 582
Db 1612 ACTACCAAAATCCAGCTGTAAAGCGAATTCATCTGGGAGCTCTCTTAAGGTTGTCAA 1671
QY 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle 602
Db 1672 GGACCTGGTCATACAGAGGGGATTAAT-----GATTTCAAGATCATTTCAAAAT 1725
QY 603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622

Db 1726 ACATGTCAACACTCAAAATTTCAACAATCGTATTATTAAGAATTCGTTATGCTTCAAT 1785
QY 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db 1786 GGAAGCGCAAAATACACGAGCTGTTATAAATCTTAGTATCCAGGGGTAGCAGAACTG--- 1842
QY 643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
Db 1843 GGTATGCACTCAACCCCACTTTTCTGTTACAGATTATACGAATTTAAATAATAAGAT 1902
QY 663 PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
Db 1903 TTTCAAGTACTAGAAATTTCTACGAGGTGAATTTGCTCCAAATCAAAACATATCTCTT 1962
QY 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleAspLysIleGluPhe 701
Db 1963 GTGTTTAATCGTTCCGATGTATATACAAACACACACAGTACTTATTGATAAATTTGAATTT 2022
QY 702 IleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGln 721
Db 2023 CTGCCAATTAATCTCGTTCTATAAGAGAGGATAGAGAAACAAATAATTAGAAACAGTCAA 2082
QY 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db 2083 CAAATTAATTAATCAATTTTATGCAATCTCTATAAAA 2118
RESULT 3
US-09-224-024-30
; Sequence 30, Application US/09224024
; Patent No. 6056953
; GENERAL INFORMATION:
; APPLICANT: Leslie Hickie
; APPLICANT: Jewel Payne
; TITLE OF INVENTION: Materials and Methods for the Control of
; TITLE OF INVENTION: Calliphoridae Pests
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/224,024
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/856,226
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA79
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2061 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-224-024-30
Alignment Scores: 1.61e-85 Length: 2061
Pred. No.:

Score:	899.50	Matches:	247
Percent Similarity:	48.32%	Conservative:	126
Best Local Similarity:	31.99%	Mismatches:	270
Query Match:	23.25%	Indels:	129
DB:	3	Gaps:	30

US-10-783-417-2 (1-735) x US-09-224-024-30 (1-2061)

QY	1	MetAsnGlnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr	20
DB	1	ATGATCCATATCAAAATAGAAATGAATATGAATATTCATGTCCATCAATGGTGT	60
QY	21	PheProAsnAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro	40
DB	61	-----AGCAAGCTAATACTATCTAGATATCCATAGCAATATAGCCAAATCAACCA	114
QY	41	LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly	60
DB	115	CTGAAAAACACGAATTTACAAAGATTGGCTCAATGTGTCAAGATAATCAACATATGGC	174
QY	61	AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaIleAlaValSerAlaGlyThrIle	80
DB	175	AATAATCGCGGGGAATTTGTAGTCTCGAACTATTGTTCGAGGTAGTCAGGTATTATT	234
QY	81	ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly	100
DB	235	GTAGTAGGAATATGTTAGGA-----GCTTTTGTGCCCT-----	270
QY	101	IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly	120
DB	271	GTCTAGTCGAGGTATAATATCTTTTGGGACITTTGTTCGGATCTTTTGG-----CAAGGA	327
QY	121	GluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr	140
DB	328	TCGTGACCCTGC AAAATGTTTGGCAGGATTTGTTAAACATCGGA-----GGAAAGG	375
QY	141	ProLeuThrGlu-----SerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPhe	157
DB	376	CCATATCAAGAAATAGATAAAACATAAATTAATGTACTTCTATCGTAACACCTATA	435
QY	158	ArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeu	177
DB	436	AAAAATCAACTTGATAAATCAAGAAATTTTCGATAAATGGGCGCCAGCACGT-----	489
QY	178	GlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIle	197
DB	490	-----ACACACGCTAATGCTAAAGCAGTACATGATCTCTTTACT	528
QY	198	ArgPheGluAsnValHisAsn-----AspPheIleArgGluIleProGlyPheGln	214
DB	529	ACCTTAGAACCTATAATAGATAAAGATTTAGATATGTTAAATAATATGCTAGCTATCGA	588
QY	215	LeuGluThrTyrLysThrLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeu	234
DB	589	ATACCAACA-----CTCCCTGCATATGCACAAATAGCTACTTGGCACTTG	633
QY	235	AsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSer	254
DB	634	AATTATTAATAACATGCTGCTACTTACTTACAAATATAG-----CTGCAAAATCAA	684
QY	255	GlnIleGluProAsnAlaGlyThrSerAspAspTyrLysLeuLeuLysGluAsn	273
DB	685	GGTATAAATCCAAAGTACTTTCAATTCATCTAATTAATCTATCAGGCTATTTAAACGTA	744
QY	274	IleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArg	293
DB	745	ATACAGAATATATCTGACTATTGTATATCAAAACCTACAACTCAGGACTTAATGATTAGA	804
QY	294	AspGluProAsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIleThr	313
DB	805	ACTAATACTAACCAACATGAAATATGATATATCTTACCGTTTAGAAATGACTCTAACT	864
QY	314	ValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIle	333

DB	865	GTGTTAGATCTTATGCTATTTTCCAAATATGACCCAGAAAATAT-----	912
QY	334	GlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThrGlu	353
DB	913	-----CCAATAGGAGTTAAATCTGAATCTACAGAGAAAGTTTAT---ACGAAT	957
QY	354	IleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyr	373
DB	958	GTTAATCATTAGCATACATT-----AGAACCATACAGAACTAGAAAAAT	999
QY	374	AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyrThr	393
DB	1000	GGATTAACTAGA---AATCCTACATTATTACTTGGATAAACCAAGGGCGTTTACACA	1056
QY	394	GluAsnThrAsn-----PheGlyAsnArgLeu	402
DB	1057	AGAAATTCGAGACATTTCTTGATGATATTTTCTTTTACAGGTAACAGATG	1116
QY	403	ValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyr	422
DB	1117	-----GCCTTTACATATACTAATGATGATCGCAACATAATCTCG	1155
QY	423	GlyGluArgThrGly-----SerProThrThrLysThrIleArgProPheGluSer	439
DB	1156	GGAGCGGTTCTAGGACATATTATTCTCAAGACATCCAAAGTATTTCTTTTATAGA	1215
QY	440	Tyr-----LysValSerIleValThrAspArgGlnSerProProValSerPro	455
DB	1216	AACAAACCTATTGATAAGTCGAAATGTCAGACATAGAGACTACAGATATAATA---	1272
QY	456	IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn	475
DB	1273	-----TATGAAATGATATTTTTCGAAATAGCATGAA	1305
QY	476	ThrLeuLysTyrSerAlaGlyGlySerLeu---SerAsnTyrGlnAsnThrThrPhe	494
DB	1306	GTATTTGATATTATCCAAATTCACATAGAAATAATATAAAGAACTGATTTCTTAT	1365
QY	495	GlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsn	514
DB	1366	ATGATTTCCAAAACAAACATCGAAAAAT-----	1392
QY	515	AsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGly	534
DB	1393	AAAGATATGTCATCTATCTATATATAAAGCTGATAATATATATTTTCAGTAGTT	1452
QY	535	LeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArg	554
DB	1453	AGAAAAGAGAGAGAGTT-----GCATTTAGTTGACACACATAGTGTGATTTC	1503
QY	555	TyrAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnLeu	574
DB	1504	CAMATCAATAGATTATAGATAACATCAACCAATCCACGCTCTAAAGCTTTGAAGGTA	1563
QY	575	AspThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeu	594
DB	1564	AGTTCTGATTCGAAATTTGAAAGGCTCTGTCACACAGGTGGAGACTTGTGTAATCTT	1623
QY	595	GlnSerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPhe	614
DB	1624	AAAGATAGTAGGATTTTAGAGTTTAGATT---TTAAAAAATGTTTCTCGCAATATCAA	1680
QY	615	IleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThr	634
DB	1681	GTACGTATTCGTTATGCTACTTAATGCT-----CCAAAGACAACTATTC	1725
QY	635	IleProGlyValIleGlyIleProGlnArgLeuAsnAsnThrPheSerGlyThrAsn	654
DB	1726	TTAACCGAATAGTACTATTAAGTGTGGAG---CTCCCTAGTACCACCTCCCGCCAAAC	1782
QY	655	TyrAsn-----AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThrVal	672

Db 1783 CCAATGCTACAGATTAAACATATGCAGATTTTGGATATGATTAACATTTTCCAAGACAGTT 1842

Qy 673 -----ThrLeuProLeuAsnArgAsnIleProPheIle 683

Db 1843 CCAATAAACAACATTGAAGGAGAGACACTTTTATTATGACC-----TTA 1887

Qy 684 PheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspIleValIleGluPheIlePro 703

Db 1888 TATGGTACACCAATCATTCATATATATA---TATATTGACAAATCGAATTTATTCCA 1944

Qy 704 IleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLys 723

Db 1945 ATCACTCAATCTGTATTAGATTATACAGAGACAAATATAGAAAAACACAGAAATA 2004

Qy 724 IleAsnThrPhePheThrAsnHisThrLysThrLeu 735

Db 2005 GTGAATGATTTATTGTTAAATTAACAAAGTTCTT 2040

RESULT 4

PCT-US94-07902-30

Sequence 30, Application PC/TUS9407902

GENERAL INFORMATION:

APPLICANT: Street address: 4980 Carroll Canyon Road

APPLICANT: City: San Diego California

APPLICANT: State/Province: California

APPLICANT: Country: US

APPLICANT: Postal code/Zip: 92121

APPLICANT: Phone number: (619) 453-8030 Fax number: (619) 453-6991

APPLICANT: Telex number:

TITLE OF INVENTION: Materials and Methods for the Control of

TITLE OF INVENTION: Calliphoridae Pests

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: David R. Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/07902

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Saliwanchik, David R.

REGISTRATION NUMBER: 31,794

REFERENCE/DOCKET NUMBER: MA79

TELECOMMUNICATION INFORMATION:

TELEPHONE: 904-375-8100

TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 2061 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

PCT-US94-07902-30

Alignment Scores:

Pred. No.: 1.61e-85 Length: 2061

Score: 899.50 Matches: 247

Percent Similarity: 48.32% Conservative: 126

Best Local Similarity: 31.99% Mismatches: 270

Query Match: 23.25% Indels: 129

DB: 6 Gaps: 30

US-10-783-417-2 (1-735) x PCT-US94-07902-30 (1-2061)

Qy 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20

Db 1 ATGAATCCATATCAAAATAAGATGAATATGAAATATTCAATGCTCCATCCAAATGGTTT 60

Qy 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40

Db 61 -----AGCAAGTCTAATAACTATTCTAGATATCCATTAGCAAAATAAGCCAAATCAACCA 114

Qy 41 LeuGlnAsnThrAsnTyrLysGluIleLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60

Db 115 CTGAATAACACAGATTAACAAGATTGGCTCAATGTGTGTCAGATAATCAACATATGGC 174

Qy 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80

Db 175 AATAATGCGGGGAATTTTGTAGTTCTGAAACTATTGTGGAGTTAGTGCAGGTATTATT 234

Qy 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100

Db 235 GTAGTAGGAACATATGTTAGGA-----GCTTTTGTCTGCCCT----- 270

Qy 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120

Db 271 GTCTTACTGCAGTATPAATATCTTTTGGACTTTGTTCGCGACTTTTGG-----CAAGGA 327

Qy 121 GluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThr 140

Db 328 TCTGACCCTGCAATGTTTGGCAGGATTTGTTAAACATCGGA-----GGAAGG 375

Qy 141 ProLeuThrGlu-----SerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPhe 157

Db 376 CCTATACAGAAATAGATAAAACATAATTAATCTACTAATCTTCTGTAAACACTATA 435

Qy 158 ArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeu 177

Db 436 AAAAAACAACCTTCATATAATCAAGAAATTTTCGATAAATGGGAGCCAGCACGT----- 489

Qy 178 GlnAlaProGlyLeuProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIle 197

Db 490 -----ACACAGCTAATGCTAAAGCAGTACATGATCTCTTTACT 528

Qy 198 ArgPheGluAsnValHisAsn-----AspPheIleArgGluIleProGlyPheGln 214

Db 529 ACCTTAGAACCTATATAAGATAAAGATTTAGATATGTTAAAAATAAATAGCTAGCTATCGA 588

Qy 215 LeuGluThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeu 234

Db 589 ATACCAACA-----CTCCCTGCATATGCACAAATAGCTACTTGGCAGCTTG 633

Qy 235 AsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSer 254

Db 634 AATTTATAAACATGCTGCTACCTATTACATATATATGG-----CTGCATAAATCAA 684

Qy 255 GlnIleGluProAsnAlaGlyThrSerAspAspTyrTyrLys-----LeuLeuLysGluAsn 273

Db 685 GGTATAAATCCAAAGTACTTTCAATTCATCTAATTAATCTATCAGGCTATTATAAAGCTAAA 744

Qy 274 IleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArg 293

Db 745 ATACAAGATATATCTGACTATTGTATACAACCGTACCAATGCAGGACTAACTATGATTAGA 804

Qy 294 AspGluProAsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIleThr 313

Db 805 ACTAATACTAACGCAACATGGAATATGTAATAATCTTACCCTTTAGAAATGACTCTAAT 864

Qy 314 ValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIle 333

Db 865 GTGTAGATCTTATTGCTATTATTTTCCAAATTAATGACCCAGCAAAATAT----- 912

Qy 334 GlyGlyIleGluValLysGlyIleLysAsnGlnLeuThrArgGluIleTyrThrGlu 353

Db 913 -----CCAATPAGGAGTTAAATCTGAACCTTACCAGAGAGTTTAT---ACGAAT 957

Db 1807 TATCTAATAAATGAATGATTTTACATATATGATTTCAATATATAGATATATCCA 1866
QY 670 SerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspVal 689
Db 1867 AGAGTCATTTTCAGTAAATGCTTCTTCAACATACAGAGGTTATCTATAGGTATACAAACG 1926
QY 690 SerAsnSerIleLeuIleAspLysIleGluPheIleProIleThrSerSerMetHis 709
Db 1927 AATACAAATTTATTTATTTAGACCGAATCGAATTCATCCAGTAGATGAGACATATGAA 1986
QY 710 GluAsnArgGluLysGlnLysLeuGluThrIleGlnThrIleAsnThrPhePheThr 729
Db 1987 GCGAAACGGAT-----TTAGAGCGCAAGAAAGCAGTGAATGCGTGTGTTTACG 2037
QY 730 Asn 730
Db 2038 AAT 2040

RESULT 6

US-08-315-468-5
; Sequence 5, Application US/08315468
; Patent No. 5554534
; GENERAL INFORMATION:
; APPLICANT: Michaels, Tracy Ellis
; APPLICANT: Fonceerrada, Luis
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Process for Controlling Scarab Pests
; TITLE OF INVENTION: with Bacillus thuringiensis Isolates
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/315,468
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,941
; FILING DATE: 01 FEB 1993
; APPLICATION NUMBER: 07/828,430
; FILING DATE: 30-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/808,316
; FILING DATE: 16-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA73.C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1953 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: tolworthi

INDIVIDUAL ISOLATES: 43F
IMMEDIATE SOURCE: E. coli XL1-Blue (pML1,98-4), NRRL B-18291
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1953
US-08-315-468-5
Alignment Scores:
Pred. No.: 1.05e-61 Length: 1953
Score: 676.00 Matches: 209
Percent Similarity: 43.76% Conservative: 117
Best Local Similarity: 28.05% Mismatches: 283
Query Match: 17.47% Indels: 136
DB: 29
US-10-783-417-2 (1-735) x US-08-315-468-5 (1-1953)
QY 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 1 ATGAATCCAAACAAT---CGAAGTGAATATGATACGNTAAAGGTTTACACCTAACAGTGAA 57
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 58 TTGCGA-----ACTAACCATATCATATCTTTAGCTGACAAATCCAATTCGACA 108
QY 41 LeuGlnAsnThrAsnTyrLysGluTrpLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
Db 109 CTAGAAGAATTAATTAATAAGAAATTTTAAAGAATGACTGCAGACAAATTTCTACG----- 162
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80
Db 163 -----GAAGTGCTAGACAGCTCTACAGTAAAGATGACAGTTGGGACAGCAATTTCT 213
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
Db 214 GTTGAGGACAGATTTAGGTGTTGTAGGG-----GTCCA 249
QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120
Db 250 TTTGCTGGGCGCTCACCTTCATTTTATCAATCTTTTAAACGCTATATGGCAAGT--- 306
QY 121 GluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db 307 -----GATGCTGACCCATGGAAGGCTTTTATGGCACAAGTGCAGAACTACTGATAGATAAG 360
QY 141 ProLeuThrGluSerIleLysGlnLysIleGlnThrLeuGluGlyPheArgGlnIle 160
Db 361 AAAATAGAGGAGTATGCTTAAAGTAAAGCTCTTCGAGAGTTACAGGGCTTCAAAATAAT 420
QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaPro 180
Db 421 TTTGAAGATTATGTAATGCGTTGGATTCTCTGGAAGAAA----- 459
QY 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGlu 200
Db 460 ---GCGCTGTAAATTTACGAAGTGAAGAACCAAGATCGAATAAGAGAACTTTTTTCT 516
QY 201 AsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThr 220
Db 517 CAAGCAGAAAGCCATTTTTCGTAATTCATGCGCTCATTTGCGGTTTCCAAATTCGAAGTT 576
QY 221 LeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGly 240
Db 577 CTGTTTCTACCAACATATGCAACAGCTCAATATACACATTTATGCTATTAAAGATGCT 636
QY 241 AlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAla 260
Db 637 CAAGTTTTTGGAGAAGATGGGATAT----- 663
QY 261 GlyThrSerAspAsp-----TyrTyrLysLeuLeuLysGluAsnIleProLysTyr 277
Db 664 ---TCTTCAGAAGATATTGCTGAATTTTATCAAGACAAATTAATAACTTACGCAACAATAC 720

ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: tolworthi
; INDIVIDUAL ISOLATE: 43f
; IMMEDIATE SOURCE:
; CLONE: E. coli XL1-Blue (pM1.98-4), NRRL B-18291
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1953
; US-07-941-650A-3

Alignment Scores:

Pred. No.: 1.05e-61 Length: 1953
Score: 676.00 Matches: 209
Percent Similarity: 43.76% Conservative: 117
Best Local Similarity: 28.05% Mismatches: 283
Query Match: 17.47% Indels: 136
DB: 3 Gaps: 29

US-10-783-417-2 (1-735) x US-07-941-650A-3 (1-1953)

QY 1 MetAsnGlnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
DB 1 ATGAATCCAAACAAT---CGAAGTGAATATGATACGATAAAGGTTACACCTTAACAGTGAA 57
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
DB 58 TTGCCA-----ACTAACCAATAATCAATATCTTTAGCTGACAAATCCAAATTCGACA 108
QY 41 LeuGlnAsnThrAsnTyrLysGluTrpLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
DB 109 CTAGAGAATAAATTAAAGAAATTTTAAAGATGACTGACAGACAATTCACG----- 162
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80
DB 163 -----GAAGTGCTAGACAGCTCTACAGTAAAGATGCAGTTCGGACAGGAATTCCT 213
QY 81 ValSerGlyThrLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
DB 214 GTTGTAGGACAGATTTTAGGTGTGTAGGG-----GTTCCA 249
QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120
DB 250 TTTGCTGGGCGCTCACTTCATTTATCAATCATTTCTTAACGCTATATGCGCAAGT--- 306
QY 121 GluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
DB 307 -----GATCGTGAACCATGAAGGCTTTTATGGCACAAAGTGAAGTACTGTATAGATAAG 360
QY 141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
DB 361 AAAATAGAGGAGTATGCTAAAGTAAAGCTCTTCGACAGATTACAGGCTCTTCAAAATAAT 420
QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaPro 180
DB 421 TTTGAAGATTATGTAATGCTGGTGGATTCTCTGGAAGAAA----- 459
QY 181 GlyLeuProProSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGlu 200
DB 460 ---GGCGCTGTAATTTACGAAGTCGAAGAACCAAGATCGAATAAGAGAACTTTTTCCT 516
QY 201 AsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThr 220
DB 517 CAAGCAGAAAGCCATTTTCGTAATTCATCGCGTCATTTGCGGTTTCCAAATTCGAAGTT 576
QY 221 LeuLeuLeuProTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGly 240
DB 577 CTGTTTCTACCAATATGACAAAGTGCAGCTGCAATACACATTTATTGCTATTAAAGATGCT 636
QY 241 AlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAla 260
DB 637 CAAGTTTTCGGAAGAAGATGGGATAT----- 663

QY 261 GlyThrSerAspAsp-----TyrTyrLysLeuLeuLysGluAsnIleProLysTyr 277
DB 664 ---TCTTCAGAGATATTTGCTGAATTTTATCAAGACAAATTAATAACTTACGCAACAATAC 720
QY 278 SerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsn 297
DB 721 ACTGACCATGCTGTCATTTGGTATAATGTTGGATTAAATAGTTTAAAGAGTTCAACTTAT 780
QY 298 MetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThr 317
DB 781 GATCATGGGTCAAATTTAAACCGTTTTCGACAGAAATGACATTAACTGCTATTAGATCTA 840
QY 318 IleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGlu 337
DB 841 ATTGTATTATCCCATTTTATGATGTTTCGGTTATACTCA----- 879
QY 338 ValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrGluIleAsnPheAsp 357
DB 880 ---AAAGGAGTTAAACAGAACTAACAGAGACATTTTACAGATCCAATT---TTTACA 933
QY 358 ArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArg 377
DB 934 CTCAATGCTCTTCAAGAGTATGACCAACTTTTTCGAGTATAGAAACTCTATTTCGAAAA 993
QY 378 AlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsn 397
DB 994 CCT-----CATTTTATTTGATTATTTCGGTGGGATTGAATTTTCATACG----- 1035
QY 398 PheGlyAsnArgLeu---ValGlyLysSerAsnArgAspAlaProThrTyr----- 413
DB 1036 -----CCTCTTCGACCTGGTTACTTCGGGAAAGATTCTTCAATTATTGGTCTGGT 1086
QY 414 -----SerAsnThrIleThrGluThrLeuTyr 422
DB 1087 AATTATGTAGAACTAGACCTAGTATAGGATCTAATGATACATCAATCCTCCCACTTTAT 1146
QY 423 GlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSer----- 439
DB 1147 GGAGATAAATCTATTGAACCTATACAAAAGCTA-----AGCTTTGTGGACAAAAGTT 1200
QY 440 TyrLysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHis 459
DB 1201 TATCGAACTATAGCTAATACAGACATAGCGGCTTTTCGGATGGCAAGATA-----TAT 1254
QY 460 PheIleIleAsnGlnIleGluLeu---TyrLeuAsnGlySerSerAsnAsnThrLeuLys 478
DB 1255 TTTGGTGTACGAAAGTTGATTATTAGTCAATATGATGATCAAAAAATGAACT----- 1308
QY 479 TyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArg 498
DB 1309 -----AGTACACAAACATATGATTTCAAAAAGATACAAATGGCTATTTAGGT 1353
QY 499 LysLysAspCysAsnLeuValIleAspProGlyCysSer-----ProAsnPheAsnAsn 516
DB 1354 GCACAGATCTTATCGACCAATTTACCACAGAAACACAGATGAACCACTTGAAAAAGCA 1413
QY 517 TyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGln 536
DB 1414 TATAGTCATCAGCTTAATTAGCGAGAATGTTCTTA-----ATGCAG 1455
QY 537 LeuGlnIleLeuAspThrGlyValLeuGlyTrpThrHisSerSerValAspArgTyrAsn 556
DB 1456 GACCGCTGCGGAACAATTTCCATTTTTCCTTTGGACACATAGAAAGTAGAGCTTTTAT 1515
QY 557 AlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThr 576
DB 1516 ACAATTGATGCTGAAAAAATTTACTCACTTCAGTAGTGAAGCATATGCTTGTCTTCA 1575
QY 577 AsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln--- 595
DB 1576 GGCGCTTCCCATTTATGAAGGTCAGGATTCACAGAGGAAATTTACTATTTCTTAAAGAA 1635
QY 596 -----SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSer 612

Db 1636 TCTAGTAATTCATGCTAAATTTAAAGTTACCTTAAATTCAGCAGCCTGTTTACACGA 1695
Qy 613 TyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSer 632
Db 1696 TATCGCGTAAGATAAGCTAGCTTCAACG-----ACTAACCTACGA 1737
Qy 633 LeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGly 652
Db 1738 CTTTCGTG-----CAAAATTCNAACAATGATTTTCTGTG 1773
Qy 653 ThrAsnTyrAsn-----AsnLeuGlnTyrGlyAspPheGlyTyr-----PheGln 667
Db 1774 ATCTACATTAATAAACTATGATATAGAT---GGTGAATTAACATATCAACATTTGAT 1830
Qy 668 PheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAla 687
Db 1831 TTGCGCACTAGT-----AATCTAATATGGGATTCCTCGGTGATACAAAT 1875
Qy 688 Asp-----ValSerAsnSerIleLeuIleAspLysIle 699
Db 1876 GACCTTTATAATAGCAGCAATCTTCTGTTCTTAATGAAAAAATCTATATAGATAAGATA 1935
Qy 700 GluPheIleProfile 704
Db 1936 GAATTTATCCCACTA 1950

RESULT 8
PCT-US92-00040-1
Sequence 1. Application PC/TUS9200040
GENERAL INFORMATION:
APPLICANT: Donovan, William P.
APPLICANT: Ruper, Mark J.
APPLICANT: Slaney, Annette C.
TITLE OF INVENTION: BACILLUS THURINGIENSIS cryIIIC(b) TOXIN
TITLE OF INVENTION: GENE AND PROTEIN TOXIC TO COLEOPTERAN INSECTS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
ADDRESSEE: Nadel
STREET: 1601 Market Street, 36th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/00040
FILING DATE: 19920103
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/649,562
FILING DATE: 31-JAN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Egolf, Christopher
REGISTRATION NUMBER: 27633
REFERENCE/DOCKET NUMBER: 7205-29 P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-757-1590
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2430 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..2099

PCT-US92-00040-1
Alignment Scores:
Pred. No.: 3,24e-60 Length: 2430
Score: 663.50 Matches: 217
Percent Similarity: 41.87% Conservative: 115
Best Local Similarity: 27.36% Mismatches: 300
Query Match: 17.15% Indels: 161
DB: 6 Gaps: 31

US-10-783-417-2 (1-735) x PCT-US92-00040-1 (1-2430)

Qy 1 MetAsnGlnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 144 ATGAATCCAAACAAT---CGAAGTGAACATGATACGATAAGGTTACACCTAACACGAA 200
Qy 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 201 TTGCCA-----ACTAACCATATCAATATTCCTTTAGCTGACAATCCAAATTCGACA 251
Qy 41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGln---GlyAsnThrGlnTyr 59
Db 252 CTAGAAGAATTAATATAAGAAATTTTAAAGATGACTGAAGACAGCTTCTACCGAAGTG 311
Qy 60 GlyAspAsnPheGluThrPheAlaSerAlaAspThrIleAla---AlaValSerAlaGly 78
Db 312 CTAGACAAC-----TCTACAGTAAAGATGCAGTTGGGACAGGA 350
Qy 79 ThrIleValSerGlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyPro 98
Db 351 ATTTCTGTTGTAGGGCAGATTTTAGGTGTGTAGGA----- 386
Qy 99 IleGlyIleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPheTyrPro 118
Db 387 GTTCCATTTGCTGGGCACCTCACTTCATTTTATCAATCACTTCTTGACACTATATGCCA 446
Qy 119 AlaGlyGluGlnAspLysThrValThrGlnPheIleLysMetGlyGluIlePheVal 138
Db 447 AGT-----GATGCTGACCCATGAGAGCTTTTATGGCACAAGTTGAAGTACTGATA 497
Qy 139 AspThrProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArg 158
Db 498 GATAAGAAATAGAGAGATATGCTAAAGTAAAGCTCTTCGAGAGTTACAGGGTCTTCAA 557
Qy 159 GlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGln 178
Db 558 AATAATTTCGAAGATTAATGTTAAATGCTGTTAAATCTCGGAAGAAA----- 602
Qy 179 AlaProGlyLeuProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArg 198
Db 603 -----ACACCTTTAAGTTTTCGGAAGTAAAGAACCAAGATCGAATAAGGGAACCTT 653
Qy 199 PheGluAsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyr 218
Db 654 TTTTCTCAAGCAGAAAAGTCAATTTTCGTAATTCATGCGCTCATTTGAGTTTCCAAATTC 713
Qy 219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAsnPheHisLeuAsnLeuGln 238
Db 714 GAAAGTGTCTGTTCTTACCACATATGCACAAGCTGCAAAATACACATTTATTGCTATTAAA 773
Qy 239 GlnGlyAlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluPro 258
Db 774 GATGCTCAAGTTTTTGGAGAAGATGGGATATTCT----- 809
Qy 259 AsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSer 278
Db 810 TCAGAAAGATGTTGCTGAAATTTTATCAGACAATTTAAACTTACG---CAACAATACACT 866
Qy 279 AsnTyrCysAlaAsnThrTyrArgThrGlyLysAsnLeuArgAspGluProAsnMet 298
Db 867 GACCATTTGTCAAATGGGTATAATGTTGGATTAAATGTTTAAGAGGTTCAACTATGAT 926
Qy 299 LysTrpSerIlePheAsnAspTyrArgTyrMetThrThrIleThrValLeuAspThrIle 318

Db 927 GCATGGGTCAAATTTAAACCGTTTTCGAGAGAAATGACTTTAACTGATTAGATCTAATT 986
Qy 319 SerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyLeuVal 338
Db 987 GTACTTTTCCATTTTATGATGTTTCGGTATATCTCA 1022
Qy 339 LysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArg 358
Db 1023 AAGGGTGTAAACAGAACTAACAGAGACATTTTACGGATCCAAAT---TTTTCACCTC 1079
Qy 359 LeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAla 378
Db 1080 AATACTCTTCAGGAGTATGACCAACTTTTGTAGTATAGAAACTCTATTTCGAAACCT 1139
Qy 379 SerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsnPhe 398
Db 1140 -----CAATTATTGATTATTATACAGGATTTGAAATTTTCATACG----- 1178
Qy 399 GlyAsnArgLeu---ValGlyIleSerAsnArgAspAlaProThrTyr----- 413
Db 1179 -----CGTCTTCACTGTTACTCTCGGAGAGATTTCTTCAATTATGCTCGGTAAAT 1232
Qy 414 -----SerAsnThrIleThrGluThrLeuTyrGly 423
Db 1233 TATGTAGAACTAGACCTAGTATAGGATCTAGTAGAGACAATTACTTCCCACTTTATGGA 1282
Qy 424 GluArgThrGlySerProThrThrThrIleArgProPheGluSer-----Tyr 440
Db 1293 GATAAATCTACTGAACCTGTACAAAAGTTA-----AGCTTTGATGACAAAAGTTTAT 1346
Qy 441 LysValSerIleValThrAspArgGlnSerProValSerProIleGlnProHisPhe 460
Db 1347 CGAACTATAGCTAATACAGACGTAGCGGCTTGGCCGAATGCCAAGATA-----TATTTT 1400
Qy 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsn 475
Db 1401 GGTGTTACGAAAGTTGATTTTAGTCAATATGATGATCAAAAATGAAACTAGTACACAA 1460
Qy 476 ThrLeuLysTyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGln 495
Db 1461 ACATATGATTTCAAAAAGAAACAATGCGCATGAGTGCACAGGATTTCTATTGACCAATTA 1520
Qy 496 PheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsn 515
Db 1521 CCACCAAGAAACAACAGAT-----GAACCACTTGAAAAA 1553
Qy 516 AsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeu 535
Db 1554 GCATATAGTATCAGCTTAATTACGCGGAATGTTCTTA-----ATG 1595
Qy 536 GlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyr 555
Db 1596 CAGGACGGTGGTGAACAATTCATTTTACTTGGACATAGAGTGTAGACTTTT 1655
Qy 556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp 575
Db 1656 AATACAAATTGATCTCGAAAAGATTACTCAACTTCCAGTAGTGAAGCATATGCTTGTCT 1715
Qy 576 ThrAsnSerLysValIleGluGlyProGlyHisThrGlyAsnLeuValTyrLeuGln 595
Db 1716 TCAGGTGCTTCCATTTATGAAGTCCAGGATTCACAGGAGAAATTTACTATTCTCTAAA 1775
Qy 596 -----SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGln 611
Db 1776 GAATCTAGTAAATCAATGCTAAATTTAAAGTTTACATTAATTCAGCAGCTGTTACAA 1835
Qy 612 SerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIle 631
Db 1836 CGATATCGTGAAGATACGCTATGCTTCTACC-----ACTAACTTA 1877
Qy 632 SerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsnThrPheSer 651

Db 1878 CGACTTTTGTG-----CAAAATTCAAACAATGATTTTATT 1913
Qy 552 GlyThrAsnTyrAsn-----AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPhe 668
Db 1914 GTCATCTACATTAAATAAAATGATGATATAGAT-----GATGATTAAACATATCAACATTT 1970
Qy 669 ProSerThrValThrLeuProLeuAsnArgAsnIleProPhe----- 682
Db 1971 GATCTCGCACTACT-----AATCTAATATGGGGTTCTCGGGTGATACGAATGAA 2021
Qy 683 -----IlePheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIleGlu 700
Db 2022 CTTATAATAGGAGCAGATCTTTCGTTTCTAATGAAAAATCTATATAGATAGATAGAA 2081
Qy 701 PheIleProIle----- 704
Db 2082 TTTATCCAGTACAAATTTGTAAGGAGATTTTGAATGTAGGCGCATGTTCAAAATGAAAGA 2141
Qy 705 -----ThrSerSerMetHisGlnAsnArgGluLys 714
Db 2142 ATAGGAAGGTGAATTTTTCATGCTTAGGAAAGATTTCTTTTAAGAAAAAGCAACATGGAAGAAG 2201
Qy 715 GlnLysLeuGluThrIleGlnThrLys---IleAsnThr 726
Db 2202 TATACAGTACAAATATTAGAAATAAAATTTATTAAACACA 2240

RESULT 9
US-08-996-441B-37
; Sequence 37, Application US/08996441B
; Patent No. 6023013
; GENERAL INFORMATION:
; APPLICANT: English, Leigh H.
; APPLICANT: Brusock, Susan M.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Bryson, James W.
; APPLICANT: Kulesza, Caroline A.
; APPLICANT: Walters, Frederick S.
; APPLICANT: Slatin, Stephen L.
; APPLICANT: Von Tersch, Michael A.
; APPLICANT: Romano, Charles
; TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996.441B
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: MECO:151
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1959 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS

297 AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAsp 316
 778 TATGATGCATGGGTCAAATTTAAACCGTTTTCGCAGAGAAATGACTTTAACTGATTAGAT 837
 317 ThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyLe 336
 838 CTAATTGTACTTTCCCATTTTATGATATTTCGGTTATCTACTCA----- 879
 337 GluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPhe 356
 880 -----ANAGGGGTAAACAGAACTTACACAGAGACATTTTACGAGTCCAAATTTTCT 933
 357 AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThr 376
 934 CTTAGGACACCA--CTTGCCTACGGACCAACTTTTTCGATAGTAAAACTCTATTTCGA 990
 377 ArgAlaSerPheLysIlePheSerPheLeuGluGlnPheIlePheTyrThr----- 393
 991 AAACCT-----CATTTATTGTATTATTACAGGGGATTGAATTTTCATACGGCTTCCAA 1044
 394 -----GluAsnThrAsnPhe-----GlyAsnArgLeuValGlyIle 405
 1045 CCTGGTTACTTTCGGAAAGATCTCTTCAATTATTGCTGGTAAATTAT-----GTA 1095
 406 SerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArg 425
 1096 GAAACTTAGACCTAGTATAGGATCTAGTAAAGACAAATTACTTCCCACTTTTATGCGAGATAA 1155
 426 ThrGlySerProThrThrIleThrIleArgProPheGluSer-----TyrLysVal 442
 1156 TCTACTGAACCTGTACAAAGCTA-----AGCTTTGATGGCAAAAAGTTTATCGAACT 1209
 443 SerIleValThrAspArgGlnSerProPro-----ValSerPro 455
 1210 ATAGCTAATACAGACTAGCGGCTCGCGCAATGGTAAGTATATTAGTGTGTTCGAAA 1269
 456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn 475
 1270 GTTGAT-----TTTAGTCAATATGATGATCAAAAATAAGAACTAGTACACAA 1317
 476 ThrLeuLysTyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGln 495
 1318 ACATATGATTCAAAAAAGAACAAATGGCCATGTAAAGTGCACAGGATTTCTATTGACCAATTA 1377
 496 PheProArgLysIleAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsn 515
 1378 CCGCCGAAACACACAGAT-----GAACCCCTGTGAAAAA 1410
 516 AsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeu 535
 1411 GCATATAGTCATCAGCTTAATTACCGGAATGTCTTCTTA-----ATC 1452
 536 GlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyr 555
 1453 CAGGACCGCTGGAAACAAATTCCTATTTTCTTGGACACATAGAAAGTGTAGACTTTTTT 1512
 556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp 575
 1513 AATCAAAATTGATGCTCGAAAGATTACTCAACTTCGATGTGAAGCATATGCTTGTCT 1572
 576 ThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln 595
 1573 TCAGGTGCTTCCATTTATGAGGTCAGGATTCACAGGAGGAAATTTACTATTCTCTAA 1632
 596 -----SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGln 611
 1633 GAATCTAGTAATTCAAATTCGTAAATTTAAAGTTTACATTAAATTCAGCAGCCTTGTGTACAA 1692
 612 SerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIle 631
 1693 CGATATCGTGAAGATAACGCTATGCTTCTTACC-----ACTAACTTA 1734
 632 SerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSer 651


```

880 -----AAGGGGTTAAACAGAACTAAACAGAGACATTTTACGGATCCAAATTTTCT 933
QY 357 AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThr 376
Db 934 CTTAGACACCA---CTTGCGTACGGACCAACTTTTTTGGATATAGAAACTCTATTTCGA 990
QY 377 AtgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr----- 393
Db 991 AAACCT-----CATTTATTGATTATTACAGGGGATTGAATTTTCATACGCGTTCCAA 1044
QY 394 -----GluAsnThrAsnPhe-----GlyAsnArgLeuValGlyIle 405
Db 1045 CCTGGTTACTTTGGGAAAGATCTTTCATATTATGTGCTCGTAATTAT-----GTA 1095
QY 406 SerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArg 425
Db 1096 GAAACTAGACCTAGTATAGGATCTAGTAGACAACTACTTCCCATTTTATGAGATATA 1155
QY 426 ThrGlySerProThrThrLysThrIleArgProPheGluSer-----TyrLysVal 442
Db 1156 TCTACTGAACCTGTACAAAGCTA-----AGCTTTGATGGACAAAGTTTATCGAACT 1209
QY 443 SerIleValThrAspArgGlnSerProPro-----ValSerPro 455
Db 1210 ATAGCTAATACAGACGTAGCGGCTTGCGCAATGGTAAGTATATTAGGTGTACGAAA 1269
QY 456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn 475
Db 1270 GTTGAT-----TTTAGTCAATATGATGATCAAAATAATGAACTAGTACACAA 1317
QY 476 ThrLeuLysTyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPheGln 495
Db 1318 ACATATGATTCAAAAAAGAAACAAATGCGCATGTAGTGCACAGCATCTTATTGACCAATTA 1377
QY 496 PheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsn 515
Db 1378 CGCCGAGAAACACACAT-----GAACCCACTGTGAAAAA 1410
QY 516 AsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeu 535
Db 1411 GCATATAGTCATCAGCTTAATACGGGAATGTTCTTA-----ATG 1452
QY 536 GlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerValAspArgTyr 555
Db 1453 CAGGACCGTGGGAACAATTCATTTTCTTGACACATAGAAAGTGTAGACTTTTT 1512
QY 556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnLeuAsp 575
Db 1513 AATACAAATTGATGCTCAAAAGATTTCAACTTCCAGTAGTGAAGCATATGCTTGTCT 1572
QY 576 ThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln 595
Db 1633 GAATCTAGTAATTCATTTGCTAAATTTAAAGTTTACATTAATTCAGCAGCTTTGTACAA 1692
QY 612 SerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIle 631
Db 1693 CGATTCGTGTGAAGTATGAGGTCCAGGATTCAGAGGAGAAATTTACTATTCTCTAAA 1734
QY 632 SerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnThrPheSer 651
Db 1735 CGACTTTTGTG-----CAAAATTCAAACATGATTTTCTT 1770
QY 652 GlyThrAsnTyrAsn-----AsnLeuGlnTyrGlyAspPheGly 664
Db 1771 GTCATCTACATTAAATAAATATGAATAAGATGATGATTATACATATCAAACTTTGAT 1830
QY 665 TyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePhe 684
Db 1831 CTCGCAACTACTAAATCTAATATATGGGGTCTTCGGGTGATAAGAATGAACCTTATAATAGGA 1890
```

```

QY 685 AsnArgAlaAspValSerAsnSerIleLeuIleLeuLeuLeuLeuLeuLeuLeuLeuLeu 704
Db 1891 GCAGAACTCTTTTCGTTCTTAATGAAAAAATCTATATAGATAGATAGATTTATCCAGTA 1950

RESULT 11
US-08-993-170A-37
; Sequence 37, Application US/08993170A
; Patent No. 6063597
; GENERAL INFORMATION:
; APPLICANT: English, Leigh H.
; APPLICANT: Brussock, Susan M.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Bryson, James W.
; APPLICANT: Kulesza, Caroline A.
; APPLICANT: Walters, Frederick S.
; APPLICANT: Slatin, Stephen L.
; APPLICANT: Von Tersch, Michael A.
; TITLE OF INVENTION: POLYPEPTIDE COMPOSITIONS TOXIC TO
; TITLE OF INVENTION: COLEOPTERAN INSECTS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,170A
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: MECO:002
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1959 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1956
; US-08-993-170A-37

Alignment Scores:
Pred. No.: 2,58e-60 Length: 1959
Score: 663.00 Matches: 208
Percent Similarity: 43.51% Conservative: 114
Best Local Similarity: 28.11% Mismatches: 292
Query Match: 17.14% Indels: 126
DB: Gaps: 27

US-10-783-417-2 (1-735) x US-08-993-170A-37 (1-1959)
QY 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 1 ATCAATCCAAACAAT---CGAAGTGAACATGAT-----ACGATAAAGGTT 42
QY 21 PheProAsnArg-----AsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsn 38
Db 43 ACACCTAAACAGTGAATGTAATGCAACCAATCAACCAATATCTTTAGCTGACAACTCCAAT 102
```

QY 39 GlnProLeuGlnAenThrAsnTyrLysGluTyrLeuAenMetCysGln---GlyAsnThr 57
Db 103 TCACACTAGAGNATTAATTAAGANATTTTAAGATGACTGAAGACAGTCTTACG 162
QY 58 GlnTyrGlyAspAenPheGluThrPheAlaSerAlaAspThrIleAla---AlaValSer 76
Db 163 GAAGTGTCTAGACAAAC-----TCTACAGTAAAGATGCGAGTTGGG 201
QY 77 AlaGlyThrIleValSerGlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSer 96
Db 202 ACAGGAATTTCTGTGTAGGCGAGATTTAGGTGTTGTAGGA----- 243
QY 97 GlyProIleGlyIleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPhe 116
Db 244 -----GTTCCATTGCTGGGCGACTCACCCTCATTTTATCATCTTCTTAACACTATA 297
QY 117 TrpProAlaGlyGluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIle 136
Db 298 TGGCCAAAGT-----GATGCTGACCCATGGAAGGCTTTTATGGCACAAGTTGNAAGTA 348
QY 137 PheValAspThrProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGly 156
Db 349 CTGATAGATAAGAAAATAGAGGAGTATGCTAAAGATAAGCTCTTGACAGAGTTACAGGGT 408
QY 157 PheArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArg 176
Db 409 CTTCAAAATTAATTTGGAAGATATGTTAAATGCGTTAAATTCCTGGAGAAGAA----- 459
QY 177 LeuGlnAlaProGlyLeuProSerSerAlaLeuGlnGlnAlaLeuThrLeuLys 196
Db 460 -----ACACCTTTAAGTTTGGGAAGTAAAGAACCCAAAGATCGNATTAAGG 504
QY 197 IleArgPheGluAenValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGlu 216
Db 505 GAACCTTTTCTCAAGCAGAAAGCATTTTTCGTAATTCATGCGCTCATTTTGCAGTTTCC 564
QY 217 ThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaPheHisLeuAsnLeu 236
Db 565 AAATTCGAAGTGTGTTTCTACCAACAATATGCAACAGCTGCAATATACACATTTATTGCTA 624
QY 237 LeuGlnGlnGlyAlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIle 256
Db 625 TTAAGAATGCTCAAGTTTTTGGGAAGAAATGGGATATCT----- 666
QY 257 GluProAsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLys 276
Db 667 -----TCAGAAGATGTTGCTGAATTTATCATAGACAAATTAACCTTACA---CAACAA 717
QY 277 TyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro 296
Db 718 TACACTGACCATTTGTGTTAATTTGTTAATGTTGATTAATGTTTAAAGAGGTTCAACT 777
QY 297 AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValIleAsp 316
Db 778 TATGATGATGATGGGTCAAAATTTAAACCGTTTTCGACAGAGAAATGACTTTTAACCTGTAATTAGAT 837
QY 317 ThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIle 336
Db 838 CTAATTTGATCTTTCCCATTTTATGATATTCGGTTATACTCA----- 879
QY 337 GluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPhe 356
Db 880 -----AAAGGGGTAAACAGAGAACTAAACAGAGACATTTTACGGATCCCAATTTTTCT 933
QY 357 AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThr 376
Db 934 CTTAGGACACCA---CTTGGCGTACGGACCAACTTTTGTAGTATAGAAAACCTTATTTCGA 990
QY 377 ArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr----- 393
Db 991 AAACCT-----CATTTATTGATTATTACAGGGGATGAATTCATACGCGCTTCAAA 1044
QY 394 -----GluAsnThrAsnPhe-----GlyAsnArgLeuValGlyIle 405

Db 1045 CCTGGTTACTTTGGGAAGATCTTTCATATTATTGCTGGTAAATAT-----GTA 1095
QY 406 SerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArg 425
Db 1096 GAAACTAGACCTAGTATAGGATCTAGTAAGACAAATTACTTCCCACTTTTATCGAGATAAA 1155
QY 426 ThrGlySerProThrThrLysThrIleArgProPheGluSer-----TyrLysVal 442
Db 1156 TCTACTGAACCTGTACAAAAGCTA-----AGCTTTGATGGACAAAAGATTATCGAACT 1209
QY 443 SerIleValThrAspArgGlnSerProPro-----ValSerPro 455
Db 1210 ATACTAATACAGACCTAGCGCTTGGCCGAATGTAAGGTATATTTAGGTGTTAGGAAA 1269
QY 456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAen 475
Db 1270 GTTCAT-----TTTAGTCAATGATGATCAAAAAAATGAAACTAGTACACAA 1317
QY 476 ThrLeuLysTyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGln 495
Db 1318 ACATATGATTTCAAAAAGAAAACAATGGCCATGTAAGTGACACAGGATTTCTATTGACCAATTA 1377
QY 496 PheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAen 515
Db 1378 CCGCCAGAAAACAACAGAT-----GAACCACTTGAAAAA 1410
QY 516 AsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeu 535
Db 1411 GCATATAGTCATCAGCTTAATACGCGGAATGTTCTTA-----ATG 1452
QY 536 GlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyr 555
Db 1453 CAGGACGCTCGTGGAAACAATTCATTTTACTTGGACACATAGAAGTGTAGACTTTTTT 1512
QY 556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp 575
Db 1513 AATCAATGATGCTGAAAGATTAACAATTCAGTAGTGAAGACATATGCCTTGTCT 1572
QY 576 ThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln 595
Db 1573 TCAGTGTCTCCATATTGAAGTCCAGGATTCACAGGAGGAAATTTACTATTCTCTAAA 1632
QY 596 -----SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGln 611
Db 1633 GAATCTAGTAATTCATTTGCTAAATTTAAAGTTACATTAATTCAGCAGCCTTGTACAA 1692
QY 612 SerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIle 631
Db 1693 CGATATCGTGTAAAGATACGCTATGCTTCTACC-----ACTAATCTTA 1734
QY 632 SerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsnAsnThrPheSer 651
Db 1735 CGACTTTTGTG-----CAAAATTCAAACATGATTTTCTT 1770
QY 652 GlyThrAsnTyrAsn-----AsnLeuGlnTyrGlyAspPheGly 664
Db 1771 GTCATCTACATTAATAAACTATGAATAAGATGATGATTTTAACATATCAACATTTGAT 1830
QY 665 TyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePhe 684
Db 1831 CTCGCAACTACTAATTTCTAATATGGGTCTCTCGGTGATAAGAAATGAACCTATAATAGGA 1890
QY 685 AsnArgAlaAspValSerAsnSerIleLeuIleAspLysIleGluPheIleProfile 704
Db 1891 CGAAGATCTTTCGTTTCTTAATGAAAAAATCTATATAGATAGATAGAAATTTATCCAGTA 1950

RESULT 12

US-08-953-775B-37

; Sequence 37, Application US/08993775B

; Patent No. 6077824

; GENERAL INFORMATION:

; APPLICANT: English, Leigh H.

/ APPLICANT: Brussock, Susan M.
 / APPLICANT: Malvar, Thomas M.
 / APPLICANT: Bryson, James W.
 / APPLICANT: Kulesza, Caroline A.
 / APPLICANT: Walters, Frederick S.
 / APPLICANT: Slatin, Stephen L.
 / APPLICANT: Von Tersch, Michael A.
 / TITLE OF INVENTION: METHODS FOR IMPROVING THE ACTIVITY OF
 / DELTA-ENDOTOXINS AGAINST INSECT PESTS
 / NUMBER OF SEQUENCES: 113
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Arnold, White & Durkee
 / STREET: P.O. Box 4433
 / CITY: Houston
 / STATE: Texas
 / COUNTRY: USA
 / ZIP: 77210
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: Patent In Release #1.0, Version #1.30
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/993,775B
 / FILING DATE: 18-DEC-1997
 / CLASSIFICATION: 514
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Kitchell, Barbara S.
 / REGISTRATION NUMBER: 33,928
 / REFERENCE/DOCKET NUMBER: MECO:150
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: 512/418-3000
 / TELEFAX: 512/474-7577
 / INFORMATION FOR SEQ ID NO: 37:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 1959 base pairs
 / TYPE: nucleic acid
 / STRANDEDNESS: single
 / TOPOLOGY: linear
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: 1..1956
 / US-08-993-775B-37

Alignment Scores:
 Pred. No.: 2,58e-60 Length: 1959
 Score: 663.00 Matches: 208
 Percent Similarity: 43.51% Conservative: 114
 Best Local Similarity: 28.11% Mismatches: 292
 Query Match: 17.14% Indels: 126
 DB: 3 Gaps: 27

US-10-783-417-2 (1-735) x US-08-993-775B-37 (1-1959)

QY 1 MetAsnGlnAsnAspAsnAsnGluTyrGluIleAspSerHisThrSerProTyr 20
 DB 1 ATGAATCCAAACAAT---CGAAGTGAACATGAT-----ACGATAAAGGTT 42
 QY 21 PheProAsnArg-----AsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsn 38
 DB 43 ACACCTTAACAGTGAATGTCACCAATCAATCAATCAATCAATCAATCAATCAATCAAT 102
 QY 39 GlnProLeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGln---GlyAsnThr 57
 DB 103 TCAACACTAGACAAATTAATTAAGAAATTTTAAAGATGATGCTGAAGACAGTTCTACG 162
 QY 58 GlnTyrGlyAspAsnPheGluThrPheAlaSerAlaAspThrIleAla---AlaValSer 76
 DB 163 GAAGTCTAGACAAAC-----TCTACACTAAAGAGATGCGATGGG 201
 QY 77 AlaGlyThrIleValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSer 96
 DB 202 ACAGGAATTTCTGTGTAGGCGCAGATTTTAGGTGTTGTAGGA----- 243

QY 97 GlyProIleGlyIleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPhe 116
 DB 244 -----GTTCCATTTGCTGGGCACTCACTTTTATCAATCAATCTTTTAACTATA 297
 QY 117 TrpProAlaGlyGluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIle 136
 DB 298 TGCCCAAGT-----GATGCTGACCCATGGGAAGGCTTTTATGGCACAAGATTGAAGTA 349
 QY 137 PheValAspThrProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGly 156
 DB 349 CTGATAGATAAGAAATAGAGGAGTATGCTAAAGTAGAAGCTCTTCGACAGATTACAGGT 408
 QY 157 PheArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrPheArgLysLeuLysArg 176
 DB 409 CTTCAAAATAAATTTTCGAAGATTTATGTTAATGCGTTAAATTCCTCGAAGAAA----- 459
 QY 177 LeuGlnAlaProGlyLeuProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLys 196
 DB 460 -----ACACCTTTTAAGTTTCGGAAGTAAAGAGCCCAAGATCGAATAAGG 504
 QY 197 IleArgPheGluAsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGlu 216
 DB 505 GAACCTTTTCTCAAGCAGAAAGTCATTTTCGTAATTCATGCGCTCATTTGCGATTTC 564
 QY 217 ThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeu 236
 DB 565 AAATTCGAAGTGTCTTTCTACCAACATATGCACAGCTGCAATATACACATTTATTGCTA 624
 QY 237 LeuGlnGlnAlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIle 256
 DB 625 TTAAGAGATGCTCAAGTTTTCGGAAGAAATGGGATATTCT----- 666
 QY 257 GluProAsnAlaGlyThrSerAspAspTyrTyrLysLeuLysGluAsnIleProLys 276
 DB 667 -----TCAGAAGATGTTGCTGAAATTTTATCATAGACAATTAACACTACA---CAACAA 717
 QY 277 TyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro 296
 DB 718 TACACTGACCATTTGTGTTAATTCGTATTAATGTTGATTAATGTTTAAATGTTTAAAGAGGTTCAACT 777
 QY 297 AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAsp 316
 DB 778 TATGATGTCATGGGTCAAAATTTTAAACCGTTTTCGACAGAGAAATGACTTTTAACTGTATAGAT 837
 QY 317 ThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIle 336
 DB 838 CTAAATTGTACTTTTCCCATTTTATGATATTTCGGTTATACTCA----- 879
 QY 337 GluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPhe 356
 DB 880 -----AAAGGGGTAAACACAGAACTAAACAAGAGACATTTTACGGATCCCAATTTTCT 933
 QY 357 AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThr 376
 DB 934 CTTTAGGACACCA---CTTGGCTACGGACCACTTTTTCGATAGTAAAGAACTCTATTTCGA 990
 QY 377 ArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr--- 393
 DB 991 AAACCT-----CATTTATTGATTTATTACAGGGGATGAAATTTACACCGCTCTTCAA 1044
 QY 394 -----GluAsnThrAsnPhe-----GlyAsnArgLeuValGlyIle 405
 DB 1045 CCTGGTTACTTTGGGAAAGATTTCTTCAATTAATTTGGTCTGTAATTAT-----GTA 1095
 QY 406 SerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArg 425
 DB 1096 GAAACTAGACTAGTAGTAGTCTAGTAGACAAATTAATCTTCCCATTTTATGGAGATATAA 1155
 QY 426 ThrGlySerProThrThrLysThrIleArgProPheGluSer-----TyrLysVal 442
 DB 1156 TCTACTGAACCTCTACAAAGACTA-----AGCTTTGATGGACAAAAGTTTATCGAACT 1209

Db 298 TGSCCAAGT-----GATGCTGACCCATGGAAGCTTTTATGGCACAAAGTTGAAGTA 348
QY 137 PheValAspThrProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGly 156
Db 349 CTGATAGATAAGAAAATAGAGAGTATGCTAAAGATAAGCTCTTGACAGAGTTACAGGGT 408
QY 157 PheArgGlnIleLeuGlnSerThrAsnThrAlaLeuAspAspTrpArgLysLeuLysArg 176
Db 409 CTTCAAAATAATTTCGAAGATTATGTTAAATCGCTTAAATTCCTGGAAGAAA----- 459
QY 177 LeuGlnAlaProGlyLeuProSerSerAlaLeuGlnGlnAlaLeuThrLeuLys 196
Db 460 -----ACACCTTTTAAGTTTCGGAAGTAAAGAACCCAGATCGAATAAGG 504
QY 197 IleArgPheGluAsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGlu 216
Db 505 GAACCTTTTCTCAAGCAGAAAGCTCATTTTCGTAATTCATGCGCGTCATTTGCAGTTTCC 564
QY 217 ThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeu 236
Db 565 AAATTTCGAAGTGTGTTTCTACCAACATATGCACAAGCTGCAAAATACACATTATTGCTA 624
QY 237 LeuGlnGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIle 256
Db 625 TTAAGAGATGCTCAAGTTTTTGGAGAAGAAATGGGGATATTCT----- 666
QY 257 GluProAsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLys 276
Db 667 -----TCAGAGATGTTGCTGAAATTTTATCATAGACAATTAACCTTACA---CAACAA 717
QY 277 TyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuLeuArgAspGluPro 296
Db 718 TACACTGACCAATGTGTAAATGCTATATGTTGGAATTAATGTTTAAAGAGGTTCAACT 777
QY 297 AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAsp 316
Db 778 TATGATGCAATGGCTCAAAATTTTACCGTTTTTCGAGAGAAATGACTTTAACTGTATTAGAT 837
QY 317 ThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIle 336
Db 838 CTAAATGTACTTTTCCATTTTATGATATTCGGTTTACTCA----- 879
QY 337 GluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPhe 356
Db 880 -----AAAGGGGTAAACAGAACTAAACAAGACATTTTACGGATCCAAATTTTCT 933
QY 357 AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThr 376
Db 934 CTTAGGACACCA---CTTGGCTGCGGACCAACTTTTTTGAGTATAGAAAACTCTATTGCA 990
QY 377 ArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr----- 393
Db 991 AAACCT-----CATTTATTTGATTAATTTACAGGGGATGAAATTCATAGCGCTCTTCAA 1044
QY 394 -----GluAsnThrAsnPhe-----GlyAsnArgLeuValGlyIle 405
Db 1045 CCTGGTTACTTTGGGAAAGATCTTTCAATTTATGCTCTGTTAATTA-----GTA 1095
QY 406 SerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyLysArg 425
Db 1096 GAAACTAGACCTAGTATAGGATCTAGTAAGACAATTTACTCCCAATTTTATGAGATATAA 1155
QY 426 ThrGlySerProThrThrLysThrIleArgProPheGluSer-----TyrLysVal 442
Db 1156 TCTACTGAACCTGTACAAAGCTA-----AGCTTTGATGGACAAAAGTTTATCGAACT 1209
QY 443 SerIleValThrAspArgGlnSerProPro-----ValSerPro 455
Db 1210 ATAGCTAATACAGACCTAGCGGCTTGGCGGAATGTAAGTATATTTAGGTGTACGAAA 1269
QY 456 IleGlnProHiePheIleleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn 475
Db 1270 GTTGAT-----TTTAGTCAATATGATGATCAAAAAAATGAAACTAGTACACAA 1317

QY 476 ThrLeuLysTyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGln 495
Db 1318 ACATATGATTCAAAAAAGAAAACAAATGGCCATGTAGTCACAGGATTTCTATTGACCAATTA 1377
QY 496 PheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsn 515
Db 1378 CCSCCAAGAAACAACAGAT-----GAAACCACTTGAAAAA 1410
QY 516 AsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeu 535
Db 1411 GCATATAGTCACTCAGCTTAATTTAGCGGAATGTTCTTA-----ATG 1452
QY 536 GlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyr 555
Db 1453 CAGGACCGTCTGGAAACAATTCATTTTACTTTGGACACATAGAGTGTAGACTTTT 1512
QY 556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp 575
Db 1513 AATACAATTTGATGCTGAAAAAGATTACTCAACTCCAGTAGTGAAGCATATGCCTTGTCT 1572
QY 576 ThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln 595
Db 1573 TCAGGTCTTCCATTTATTGAAGTCCAGGATTCACAGAGAAATTTACTATTCTTAATA 1632
QY 596 -----SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGln 611
Db 1633 GAATCTAGTAATTTCAATTCGTAATTTAAAGTTTACATTAAATTCAGCGCCTTGTCAAA 1692
QY 612 SerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIle 631
Db 1693 CGATCTCGTGAAGAATACGCTATGCTTCTACC-----ACTAACTTA 1734
QY 632 SerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsnAsnThrPheSer 651
Db 1735 CGACTTTTGTG-----CAAAATTCAAACAATGATTTTCTT 1770
QY 652 GlyThrAsnTyrAsn-----AsnLeuGlnTyrGlyAspPheGly 664
Db 1771 GTCATCTACATTAATAAACTATGAATAAGATGATGATTTAAACATATCAACATTTGAT 1830
QY 665 TyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePhe 684
Db 1831 CTGCAACTACTAATTTCTAATATGGGTTCTCGGTTGATAAGAATGAACCTTATAATAGGA 1890
QY 685 AsnArgAlaAspValSerAsnSerIleLeuIleAspLysIleGluPheIleProIle 704
Db 1891 GCAGAACTCTTCTGTTCTTAATGAATAAAATCTATATAGATAAGATGAATTTATCCAGTA 1950

RESULT 14

US-09-427-769-37

; Sequence 37, Application US/09427769

; Patent No. 6642030

; GENERAL INFORMATION:

; APPLICANT: English, Leigh H.

; APPLICANT: Brussock, Susan M.

; APPLICANT: Malvar, Thomas M.

; APPLICANT: Bryson, James W.

; APPLICANT: Kulesza, Caroline A.

; APPLICANT: Walters, Frederick S.

; APPLICANT: Slatin, Stephen L.

; APPLICANT: Von Tersch, Michael A.

; APPLICANT: Romano, Charles

; TITLE OF INVENTION: NUCLEIC ACID SEGMENTS ENCODING MODIFIED

; TITLE OF INVENTION: COLIPOPTERAN-TOXIC CRYSTAL PROTEINS

; NUMBER OF SEQUENCES: 113

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Arnold, White & Durkee

; STREET: P.O. Box 4433

; CITY: Houston

; STATE: Texas

; COUNTRY: USA

; ZIP: 77210

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/427,769
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/993,722
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Kitchell, Barbara S.
 REGISTRATION NUMBER: 33,928
 REFERENCE/DOCKET NUMBER: MECO:149
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 512/418-3106
 TELEFAX: 512/474-7577
 INFORMATION FOR SEQ ID NO: 37:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1959 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1956
 US-09-427-769-37

Alignment Scores:
 Pred. No.: 2,58e-60 Length: 1959
 Score: 663.00 Matches: 208
 Percent Similarity: 43.51% Conservative: 114
 Best Local Similarity: 28.11% Mismatches: 292
 Query Match: 17.14% Indels: 126
 DB: 3 Gaps: 27

US-10-783-417-2 (1-735) x US-09-427-769-37 (1-1959)

```

QY 1 MetAsnGlnAsnAspAsnAsnGluTyrGluIleAspSerHisThrSerProTyr 20
DB 1 ATGAATCCAAACAT---CGAAGTGAACATGAT-----ACGATAAAGGTT 42

QY 21 PheProAsnArg-----AsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsn 38
DB 43 ACACCTAACAGTGAATGTCACAACTAACCAATCAATATCTTTAGCTGACATCCAAAT 102

QY 39 GlnProLeuGlnAsnThrAsnTyrLysGluTrpLeuAsnMetCysGln---GlyAsnThr 57
DB 103 TCAACACTAGAGAAGATTAATTAAGAAATTTTAAAGATGACTGAAGACAGTCTCTACG 162

QY 58 GlnTyrGlyAspAsnPheGluThrPheAlaSerAlaAspThrIleAla---AlaValSer 76
DB 163 GAAGTCTAGAACAC-----TCTACAGTAAAGATGACAGTGGG 201

QY 77 AlaGlyThrIleValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSer 96
DB 202 ACAGGAATTCGTGTAGGCGAGATTTTAGGTGTGTAGGA----- 243

QY 97 GlyProIleGlyIleLeuAlaIleIleSerPheGlyThrLeuIleThrValPhe 116
DB 244 -----GTTCCATTGCTGGGCACTCACTTCATTTTATCAATCATTTCTTAACACTATA 297

QY 117 TrpProAlaGlyGluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIle 136
DB 298 TGCCCAAGT-----GATGCTGACCCCATGGAAGCCTTTTATGGCACAAGTTGAAGTA 348

QY 137 PheValAspThrProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGly 156
DB 349 CTGATAGATAAGAAAATAGAGGAGTAGCTAAAGTAAGTAAAGTCTTCGACAGATTACAGGGT 408

QY 157 PheArgGlnIleLeuGlnInsertyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArg 176

```

```

DB 409 CTTCAAAATAATTTTCAAGATTATGTTAATGCGTTAAATCTCTGGAAGAA----- 459
QY 177 LeuGlnAlaProGlyLeuProProSerAlaLeuGlnGlnAlaLeuThrLeuLys 196
DB 460 -----ACACCTTTAAGTTTGCAGAGTAAAGAAAGCAAGCAAGTCAATAAGG 504
QY 197 IleArgPheGluAsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGlu 216
DB 505 GAACCTTTTCTCAAGCAGAAAGTCATTTTCTGTAATTCATCCGTCATTTGTCAGTTTCC 564
QY 217 ThrTyrLysThrLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeu 236
DB 565 AAATTCGAAGTCTGTTTCTTCCACATATGCAAGCTGCAAAATACACATTTATTGCTA 624
QY 237 LeuGlnGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIle 256
DB 625 TTAAGAGATGCTCAAGTTTTCGAGAGAAAGTGGGATATTCI----- 666
QY 257 GluProAsnAlaGlyThrSerAspAspTyrTyrLysLeuLysGluAsnIleProLys 276
DB 667 -----TCAGAGATGTTGCTGAATTTTATCATAGACAAATTAATACTTACA---CAACAA 717
QY 277 TyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro 296
DB 718 TACACTGACCAATGTTGTTAATTTGGTATTAATGTTGAATTAATGTTTAAAGAGTTCAACT 777
QY 297 AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAsp 316
DB 778 TATGATGATGGGTCAAATTTAACCGTTTTTCGACAGAAATGACCTTAATCTGATTAGAT 837
QY 317 ThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIle 336
DB 838 CTAATTTGATCTTTTCCCATTTTATGATATTCGTTATATCTCA----- 879
QY 337 GluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPhe 356
DB 880 -----AAAGGGGTAAACAGAACTAACACAGACATTTTACGGATCCCAATTTTCT 933
QY 357 AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThr 376
DB 934 CTTAGGACACCA---CTTGCCTACGGACCACTTTTGTAGTATAGAAACCTCTATTCCA 990
QY 377 ArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr----- 393
DB 991 AAACCT-----CATTTATTGATTATTACAGGGGATTGAATTTTATCATACGGTCTTCAA 1044
QY 394 -----GluAsnThrAsnPhe-----GlyAsnArgLeuValGlyIle 405
DB 1045 CTTGTTTACTTTGGGAAAGATTCTTTCATTTATTGCTGCTGGTAATTTAT-----GTA 1095
QY 406 SerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArg 425
DB 1096 GAAACTAGACCTAGTATAGGATCTAGTAAGACAAATTACTTCCCAATTTTATGGAGATAAA 1155
QY 426 ThrGlySerProThrThrLysThrIleArgProPheGluSer-----TyrLysVal 442
DB 1156 TCTACTGAACCTGTACAAAAGCTA-----AGCTTGTAGTGCACAAAAGTTTATCGAACT 1209
QY 443 SerIleValThrAspArgGlnSerProPro-----ValSerPro 455
DB 1210 ATAGCTAATACAGACGTAGCGGCTTGGCCGAATGGTAAGGTATATTTAGGTGTTACGAAA 1269
QY 456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn 475
DB 1270 GTTGAT-----TTTAGTCAATATGATGATCAAAAATAATGAAACTAGTACACAA 1317
QY 476 ThrLeuLysTyrSerAlaGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGln 495
DB 1318 ACATATGATCAAAAAGAAACAATGGCCATGTAAGTGACAGGATTTCTATTGACCAATTA 1377
QY 496 PheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsn 515

```

```
Db 1378 CCGCCAGAACACAGAT-----GAACACTTGAAAAA 1410
Qy 516 AsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSertyrValIleGlyLeu 535
Db 1411 GCATATAGTCATCAGCTTAATACGGCGAATGTTCTTA-----ATG 1452
Qy 536 GlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyr 555
Db 1453 CAGGACCGTGTGGAAACAATTCATTTTACCTGGACACATGAAGTGAGACTTTTTT 1512
Qy 556 AsnAlaIleSerAspIleIleThrMetIleProAlaIleLeuGlyGlyAsnAsnLeuAsp 575
Db 1513 AATACATTTGATGCTGGAAGATATCTCAACTCCAGTAGTAGAAGCATATGCTTGTCT 1572
Qy 576 ThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln 595
Db 1573 TCAGGTGCTTCCATTTATGAAGTCCAGGATTCACAGGAGGAAATTTACTATTCTTAA 1632
Qy 596 -----SerGlnGlyArgLeuGluLeuThrCysGluThrProAsnSerThrGln 611
Db 1633 GAATCTAGTAATTCATTTGCTAAATTTAAAGTTACATTAATTCAGCAGCCTTGTTACAA 1692
Qy 612 SerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIle 631
Db 1693 CGATATCGTGAAGATACGCTATGCTTCTACC-----ACTAACTTA 1734
Qy 632 SerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsnAsnThrPheSer 651
Db 1735 CGACTTTTGTG-----CAAAATTCAAACAATGATTTCTT 1770
Qy 652 GlyThrAsnTyrAsn-----AsnLeuGlnTyrGlyAspPheGly 664
Db 1771 GTCATCTACATTAATAAACTATGAATAAAGATGATGATTTAAACATATCAACATTTGAT 1830
Qy 665 TyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIlePropheIlePhe 684
Db 1831 CTGCGAACTACTAATTTCTAATATGGGGTCTCGGGTGATAAGAATGAACATTATAATAGGA 1890
Qy 685 AsnArgAlaAspValSerAsnSerIleLeuIleAspLysIleGluPheIleProIle 704
Db 1891 GCAGAACTCTTCTGTTTCTTAATGAANAATCTATATAGATAAGATAGAAATTTATCCCA 1950

RESULT 15
US-07-973-320-3
; Sequence 3, Application US/07973320
; Patent No. 5286486
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel M.
; APPLICANT: Fu, Jenny M.
; TITLE OF INVENTION: No. 5286486el Bacillus thuringiensis Gene
; TITLE OF INVENTION: Encoding a Coleopteran-Active Toxin
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/973,320
; FILING DATE: 19921106
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/788,638
; FILING DATE: 6-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
```

```
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA68.C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3414 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: kumamotoensis
; INDIVIDUAL ISOLATE: HD867
; IMMEDIATE SOURCE:
; LIBRARY: Lamdagem (TM)-11 library of J.M. Fu
; CLONE: 867
; US-07-973-320-3

Alignment Scores:
Pred. No.: 1.18e-59 Length: 3414
Score: 660.50 Matches: 223
Percent Similarity: 43.06% Conservative: 112
Best Local Similarity: 28.66% Mismatches: 286
Query Match: 17.07% Indels: 157
DB: 2 Gaps: 37

US-10-783-417-2 (1-735) x US-07-973-320-3 (1-3414)
Qy 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 1 ATGAATTTAATAATTTAGTGGA-----TATGAA-----GATAGT----- 36
Qy 21 PheProAsnArgAsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGln 39
Db 37 -----AATAGAACATTAATAATTCCTCAATTTATCTCAATAAAGCATTTATCACC 90
Qy 40 ProLeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyr 59
Db 91 TCATTAAGAATAATGAATCACTACCAGGATTTTATCTATACTAG----- 135
Qy 60 GlyAspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThr 79
Db 136 AGGGAACAGCGCTGAAGCAGCTCGCTAGTGGTAAATACA-----GCTATTAAATCTAGT 189
Qy 80 IleValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIle 99
Db 190 AGTGTACAGGGGCTACACTAAGTCGCTAGT-----GTCCAGGT----- 231
Qy 100 GlyIleIleGlyAlaIleIleSerPheGlyThrLeuIleThr---ValPheTrpPro 118
Db 232 -----GCAAGTTTTATCACTAACTTTTACCTGAAATATACAGGCTTTTATGGCCA 282
Qy 119 AlaGlyGluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheVal 138
Db 283 -----CACGATAAAAAATATTGGGATGAATTTATGACAGAGTAGAACAACCTTATT 333
Qy 139 AspThrProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGlyGlyPheArg 158
Db 334 GAACAAAAATAGACAATAATGCAAGGAATAAAGCATCTTCAGAAATAGAGGGATTAGGA 393
Qy 159 GlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGln 178
Db 394 AATAACTTAACGATATATCAACAGGCACTTGAAGATTGG-----CTGAAC 438
Qy 179 AlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIle--- 197
Db 439 AATCCTGATGATCCA-----GCAACTATAACACGAGTATAGAT 477
```

1399	Db		TCAACTCCGGATTAT		-----GATAATGCTACTATCCCG	1431
544	Qy	ValleuGlyTrpThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIle	563			
1432	Db	ATCTTTTCTTGGACGCATAGAAAGTCGGAGTATTACAATAGAAATCTATCCAAACAAAATC	1491			
564	Qy	ThrMetIleProhlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGluGly	593			
1492	Db	ACAAAATTCACACTGTAAAAATGTATAACTAGGTGATACATCTACAGTTGTCAAAAGG	1551			
584	Qy	ProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIleThr	603			
1552	Db	CCTGGATTTACAGCTGGAGATTTAGTTAAGAGAGGGAGTAATCGTTATATAGGAGATATA	1611			
604	Qy	CysGluThrProAsnSer-----ThrGlnSerTyrPheIleArgLeuArgTyrAlaThr	621			
1612	Db	AAGGCTACCGTAACACTACACACTCTTCTCAAAATATTTCGTGTAGATTTCGATACGGCACT	1671			
622	Qy	AsnGlyValaGly-----AsnThrLeuProAsnIleSerLeuThrIleProGlyValIle	639			
1672	Db	AATGTTTCTGGACAATTCAACGGTGTTATTAATGATAAAATAACGCTT-----	1719			
640	Qy	GlyIleProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyr---AsnAsnLeu	658			
1720	Db	-----CAAAAGAAGTTTCAAAATCTGTAGAAACAATAGGTGAAGGAAAGAAGATT	1770			
659	Qy	GlnTyrGlyAspPheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArg	678			
1771	Db	ACCTATGGTTCAITTTGGATATATAGAAATATCTACGACCACTTCAATTCG-----GATAAG	1827			
679	Qy	AsnIleProPheIlePheAsnArgAlaAppValSer---AsnSerIleLeuIleIleAsp	697			
1828	Db	CATCCAAAAATCACTCTCTATTATTAAGTGTGATTTTCAGTAACAATTCATCATTTTATGTAGAT	1887			
698	Qy	LysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeu	717			
1888	Db	TCATCGAATTTATCCTGTAGATGTAAATTATGAT-----GAAAAGAAGAAAACCTA	1938			
718	Qy	GluThrIleGlnThrLysIleAsnThrPhePheThrAsnHisThrLysThrLeu	735			
1939	Db	GA AAAAGCAGAAAAGCCCGGAATACCTTGTTTTCACAGAGGAGAAAGTGCATCT	1992			

Search completed: December 11, 2005, 20:13:18
Job time : 408 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 11, 2005, 17:40:06 ; Search time 1443 Seconds
(without alignments)

4212.053 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869

Sequence: 1 MQQNDNNEIIDSHTSPY.....KLETIQKINTFTNHKTL 735

Scoring table: BLOSUM62

Xgapop	10.0	Xgapext	0.5
Ygapop	10.0	Ygapext	0.5
Fgapop	6.0	Fgapext	7.0
Delop	6.0	Delext	7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10783417/runat_01122005_141955_24239/app_query.fasta_1.903
-DB=Published Applications NA.Main -OPMT=fastap -SUFFIX=rnpbm -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62
-TRANS=human40.cdi -LIST=45 -DLOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10783417 @CGN 1 1 1026 @runat_01122005_141955_24239
-NCPU=6 -ICPU=3 -NO MMAP -LARGESUBSTR -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA.Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3869	100.0	2208	8 US-10-783-417-1	Sequence 1, Appli
2	3314	85.7	2235	8 US-10-782-570-1	Sequence 1, Appli
3	3104.5	80.2	2085	8 US-10-782-570-3	Sequence 3, Appli
4	849	21.9	4896	3 US-09-756-526A-3	Sequence 3, Appli
5	849	21.9	4896	6 US-10-345-020-3	Sequence 3, Appli
6	849	21.9	4896	6 US-10-342-821-3	Sequence 3, Appli
7	759	19.6	3684	9 US-10-929-754-2	Sequence 2, Appli
8	705	18.2	4391	9 US-10-929-754-4	Sequence 4, Appli

9	703	18.2	3633	5	US-10-032-717-3	Sequence 3, Appli
10	703	18.2	3633	6	US-10-414-637-3	Sequence 3, Appli
11	703	18.2	3633	7	US-10-606-320-3	Sequence 3, Appli
12	703	18.2	3633	8	US-10-746-914-3	Sequence 3, Appli
13	703	18.2	6613	5	US-10-032-717-28	Sequence 28, Appli
14	703	18.2	6613	6	US-10-414-637-28	Sequence 28, Appli
15	703	18.2	6613	7	US-10-606-320-18	Sequence 18, Appli
16	703	18.2	6613	8	US-10-746-914-18	Sequence 18, Appli
17	689.5	17.8	4188	5	US-10-120-544A-5	Sequence 5, Appli
18	689.5	17.8	4188	10	US-11-091-654-5	Sequence 5, Appli
19	678.5	17.5	2003	5	US-10-032-717-7	Sequence 7, Appli
20	678.5	17.5	2003	6	US-10-414-637-7	Sequence 7, Appli
21	678.5	17.5	2022	5	US-10-032-717-17	Sequence 17, Appli
22	678.5	17.5	2022	6	US-10-414-637-17	Sequence 17, Appli
23	678.5	17.5	2022	7	US-10-606-320-13	Sequence 13, Appli
24	678.5	17.5	2022	8	US-10-746-914-13	Sequence 13, Appli
25	675	17.4	3621	5	US-10-032-717-1	Sequence 1, Appli
26	675	17.4	3621	6	US-10-414-637-1	Sequence 1, Appli
27	675	17.4	3621	7	US-10-606-320-1	Sequence 1, Appli
28	675	17.4	3621	8	US-10-746-914-1	Sequence 1, Appli
29	675	17.4	3621	10	US-11-021-115-5	Sequence 5, Appli
30	675	17.4	4874	5	US-10-032-717-27	Sequence 27, Appli
31	675	17.4	4874	6	US-10-414-637-27	Sequence 27, Appli
32	675	17.4	4874	7	US-10-606-320-17	Sequence 17, Appli
33	675	17.4	4874	8	US-10-746-914-17	Sequence 17, Appli
34	669	17.3	4359	5	US-10-120-544A-3	Sequence 3, Appli
35	669	17.3	4359	10	US-11-091-654-3	Sequence 3, Appli
36	668	17.3	2088	10	US-11-018-615-5	Sequence 5, Appli
37	663	17.1	1959	7	US-10-614-076-37	Sequence 37, Appli
38	663	17.1	2358	10	US-11-018-615-28	Sequence 28, Appli
39	663	17.1	3510	10	US-11-018-615-24	Sequence 24, Appli
40	660	17.1	1959	5	US-10-232-665-3	Sequence 3, Appli
41	658	17.0	1959	7	US-10-614-076-27	Sequence 27, Appli
42	658	17.0	1959	7	US-10-614-076-35	Sequence 35, Appli
43	657.5	17.0	2025	7	US-10-606-320-73	Sequence 73, Appli
44	657.5	17.0	2025	7	US-10-606-320-79	Sequence 79, Appli
45	657.5	17.0	2025	8	US-10-746-914-73	Sequence 73, Appli

ALIGNMENTS

RESULT 1
US-10-783-417-1
; Sequence 1, Application US/10783417
; Publication No. US20040216186A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Kozziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-006, A Delta-Endotoxin Gene and
; FILE REFERENCE: 045600/274146
; CURRENT APPLICATION NUMBER: US/10/783,417
; PRIOR FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: 60/448,806
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2208
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2208)
US-10-783-417-1
Alignment Scores:
Pred. No.: 0 Length: 2208
Score: 3869.00 Matches: 735
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	8	Gaps:	0
US-10-783-417-2 (1-735) x US-10-783-417-1 (1-2208)			
QY	1	MetAsnGlnAsnAspAsnAsnGluTyrGluIleAspSerHisThrSerProTyr	20
DB	1	ATGAATCAAAATACGATATATACGAATATGAATATGAATTCGTCACCTTAT	60
QY	21	PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro	40
DB	61	TTTCCGAAACAGAAACAGTATGATTTCTAGTACCTTACAAATATCCAAATCAACCA	120
QY	41	LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly	60
DB	121	TTACAAACACAAATTTACAAAGAGTGGCTCAATATGTGTCAAGGGAATACACAATATGGT	180
QY	61	AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIle	80
DB	181	GATAATTTTCGAGACATTTGCTAGTGTGATACAAATTTGCTGAGTGTAGTACCTTAT	240
QY	81	ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly	100
DB	241	GTATCCGGTACTCTGTAGCCGTATAGGTGGCTCACTTCTATATCCGACCGATAGGA	300
QY	101	IleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGly	120
DB	301	ATAATAGTGTCTATATATATCTTTTGGTACCTTAATCACTGCTTTTGGCCGGGGA	360
QY	121	GluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr	140
DB	361	GAACAAGACAAAACATGATGGACACAAATTTATTAATAATGGAGAAATTTTGTGTGATACA	420
QY	141	ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle	160
DB	421	CCGTTAAACAGAAAGCATAAACAGCTTAAAGTTTACAAACTTTTAGAAGGATTTAGACAAATA	480
QY	161	LeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaPro	180
DB	481	TTACAAGCTAATAACAGCATATGATGATGGAGAAATTTAAAGACCTACAAAGCTCCT	540
QY	181	GlyLeuProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGlu	200
DB	541	GGATTACCAACCATCATCAGCATTTTACAAAGCTGCTTCACTCTTAAATACGATTTAG	600
QY	201	AsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThr	220
DB	601	AAATGTTTCAATGATTTTATTCGAGAAATACCTGGTTTCCAACTTGAAACTTATAAACC	660
QY	221	LeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGly	240
DB	661	CTATTACTACCTATTATGGCAAGCTGCTTAATTTTCAATTTTATTAACCAAGGT	720
QY	241	AlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAla	260
DB	721	GCTGAATGGCTGATGAATGAATGATGATATACATCTTACAAATTTGAACCTATGCT	780
QY	261	GlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyr	280
DB	781	GGAAATCATGATGACTATTATTAACCTTTTAAAGAAAAATATACCTAAATATAGTAACAT	840
QY	281	CysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTyr	300
DB	841	TGTGCAAAATACCTTATAGAACAGGACTTAAATAATCTTTAGAGACGAAACCAATATGAATGG	900
QY	301	SerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGln	320
DB	901	AGTATATTTAATGACTATCGAAGATATATGACCATTTACTGTATTAGATACCATCTCTCAA	960
QY	321	PheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyIleGluValLysGly	340
DB	961	TTTTTCTTTATATGATATAAAAAAGATATAGAGATTCAATAGGAGCAATAGAAGTAAAGGC	1020
QY	341	IleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuPro	360
DB	1021	ATTAAGAATGAACCTCACAGAGAAATTTATACAACTGAAATATAATTTGATCGTCTTCT	1080
QY	361	GlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPhe	380
DB	1081	CAACTTAGAGTTCAACCAATCTAGCTACGATGGAATATAATTTAACACACGGCAAGTTTT	1140
QY	381	LysLeuPheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsn	400
DB	1141	AAATATTTTCAATTTTAGAACAAATTTATTTTATACAGAAATATACAAATTTGGGAAT	1200
QY	401	ArgLeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr	420
DB	1201	CGTTTAGTGGTATTTCTAATCGTATCGTACCTACTTATAGCAATATCTATTAACCTGAACT	1260
QY	421	LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr	440
DB	1261	TTATATGGAGAAAGACAGGTTTCAACCAACAAAAACAAATAGACCATTTGAAATCTTAT	1320
QY	441	LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe	460
DB	1321	AAAGTTTCAATTTGTAATGATAGCAATCACTCTCTGTTTCCCTTATTCACCAACACATTT	1380
QY	461	IleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsnThrLeuLysTyrSer	480
DB	1381	ATAATTAATCAAAATTTGAACCTTTATTTAAATGGCTCATCTAACACACACATCAATATTC	1440
QY	481	AlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArgLysLys	500
DB	1441	GCAGGAGGGTCTTTATCTAATTTATCAAAACACACACTTTTTCATTTTCTTAGAAAAA	1500
QY	501	AspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIle	520
DB	1501	GACTGCAATCTAGTTATTTGATCCAGGTGTTTCCAAACTTTTAAATACTATAGTCATATT	1560
QY	521	LeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeu	540
DB	1561	TTATCCCAATTTTCAATTTATTTACTTATTTCTATGTTGATTTGATAGCTACCAATATTA	1620
QY	541	AspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerAsp	560
DB	1621	GATACAGGTATTTAGGATGGACACACAGTAGTGTGATGATATATAATCAATATCAGAT	1680
QY	561	LysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysVal	580
DB	1681	AAAAATTAACAATGATCCAGCAATCAAGGTAACCAATCTTGATACAACTCTAAGGTA	1740
QY	581	IleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeu	600
DB	1741	ATTGAAGGACCTGGTCATACAGGAGGAACTTGGTTTATTTACAAAGTCAAGGGCGTTTA	1800
QY	601	GluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAla	620
DB	1801	GAATTTACATGTGAAACTCTCTAATTTACACAATCTTATTTTATTAGCTTCGATATGCT	1860
QY	621	ThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGly	640
DB	1861	ACAAATTTGGTCTCGAAATATCTTCTCTTAATATATCTCTTCAATACCAGGAGTAAATAGGA	1920
QY	641	IleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyr	660
DB	1921	ATACCACTCAACGACTCAACAAACACTTTTCTGGTACAAATATATAATTTATTAATATAC	1980
QY	661	GlyAspPheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIle	680
DB	1981	GGAGATTTTGGTATTTTCCAAATTTCCAAATACAGTAAACATTTACCTTTAAATCGAAACATA	2040
QY	681	ProPheIlePheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIleGlu	700
DB	2041	CCATTTTATATTTAATCGTGAGATGATCAAAATTTCAATTTTAAATCATTTGATTAATTTGAA	2100

Qy	701	Phe	Thr	Leu	Thr	Leu	Ser	Ser	Met	His	Gln	Asn	Arg	Gln	Leu	Glu	Thr	Ile	720
Db	2101	TTT	ATC	CAAA	TTT	ACT	TCT	CTC	TAT	GCA	CAA	AAATAG	AGAAAA	CAAA	AAAT	TAG	AAAC	TATC	2160
Qy	721	Gln	Thr	Lys	Leu	Asn	Thr	Phe	Phe	Thr	Asn	His	Thr	Lys	Thr	Leu		735	
Db	2161	CAAA	CAAAA	TAA	TAA	TAC	ATTTT	CAC	AAATCAT	CATA	CAAAAA	CAC	TTT					2205	

RESULT 2

```

US-10-782-570-1
; Sequence 1, Application US/10782570
; Publication No. US20040210965A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-007, A Delta-Endotoxin Gene and
; TITLE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/274144
; CURRENT APPLICATION NUMBER: US/10782,570
; CURRENT FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: 60/448,812
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2235
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2235)
US-10-782-570-1

```

Qy	141	ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle	160
Db	418	CCGTTAAcAGAAGCAATAAAACAGCTAAAGTTACAAACTTTTAGAAGAGATTTAGACAAATA	477
Qy	161	LeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgIysLeuLysAcGLeuGlnAlaPro	180
Db	478	TTACAAAGCTATTATAACGATTAGATGTGGAGAAAAATTAAAGAGACTACAAGCTCCT	537
Qy	181	GlyLeuProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGlu	200
Db	538	GGATTACCACCATCATCAGCATTAACAACAGCTGCCTTGACTCTTAATAATACGATTTGAG	597
Qy	201	AsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThr	220
Db	598	AATGTTCCAAATGATTTTATTCGAGAAATACCTGGTTTCCAACTTGAAACTTATAAAACG	657
Qy	221	LeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGly	240
Db	658	CTATTACTCTATTATGCGCAGCTGCTAAATTTTCATTTAAATTTATTATCAACAAGGT	711
Qy	241	AlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAla	260
Db	718	GCTGAATTCGCTGATGAATCGAATGCAGATATACATCTTCACAAATTTGAACCTAATGCT	777
Qy	261	GlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyr	280
Db	778	GGAAACATCAGATGACTATTATTAACCTTTTAAAGAAAAATATACCTAAATATAGTAACTAT	837
Qy	281	CysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTyr	300
Db	838	TGTGCAAAATACCTATAGAGAAGCACTAAATTAACCTTCGAAACGAACCTAATATGAGATGG	897
Qy	301	SerIlePheAsnAspTyrArgTyrMetThrIleThrValLeuAspThrIleSerGln	320
Db	898	AGTATATTTAATGATATCGAGATATAGCATTTACTGTATTAGATACATCTGCTCAA	957
Qy	321	PheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluValLysGly	340
Db	958	TTTTCTTTTATGATATAAGAGATACAAAGATTCAAATAGCA-----AGAATAGGTGGC	1011
Qy	341	IleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuPro	360
Db	1012	ATTAACAACTCAACTTACACAGAGAAATTTACAACTGAAATAAATTTTGACCGCTTACT	1071
Qy	361	GlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPhe	380
Db	1072	TACCTTGAAATTCACCCCAATCTCGCTATTAATGGAATATAAATTTAAACAGCTTGAGGCTT	1131
Qy	381	LysLeuPheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsn	400
Db	1132	AGATTATTTCAATTTTAGATGAACCTATATATTTTATACAAAAATGAAACGTACGGGAAT	1191
Qy	401	ArgLeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr	420
Db	1192	CGTTTAGTTGGTATTGCGAATCGTAATAGATCTACTTATGCTACGACAGGAACTGAAATT	1251
Qy	421	LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr	440
Db	1252	ATATATGGAGAAAGACAGGTCACCCCAACAAAAACTTTTATACCAATTTGAATTCCTAT	1311
Qy	441	LysValSerIleValThrAspArgGlnSerProProValSerProIle---GlnProHis	459
Db	1312	AAAGTTTCATTTGTAACCTGATAGCAAGTAACCTCTACTTCCCTTTTCTTAACATATAC	1371
Qy	460	PheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnThrLeuLysTyr	479
Db	1372	TTTTACAATTAATCAAATTTGAACCTTTATTTAAATAAATTCACCTAGTATAATAATTAACATAT	1431
Qy	480	SerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArgLys	499
Db	1432	TCAGCTGGGGGAATTTATCTAATGATAAAAAAACCACTGATTTTCAATTTCTGTGAAAA	1491

Db 1369 TGCTCAAAATTTATGAAACACGAGAGACCTTTTAAATCAAAATTTCTACAGCTGTA 1428
Qy 147 LysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGlnSerTyrAsnThr 166
Db 1429 AAGAAATAGCATAGCTCATCTAAATGGTTTAAAGATGTATTAAACGTACTATGAAGA 1488
Qy 167 AlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeuProProSerSer 186
Db 1489 GCATTTAATGATGGAGAGA-----AATCCAAGTGCA 1521
Qy 187 AlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPhe 206
Db 1522 -----AATAGTCCAGATTGGTATCACAGAGATTGAAAACGCTCATTTCAATTTT 1572
Qy 207 IleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeuProIleTyr 226
Db 1573 GTAAGCAATATGCAAACTCCAACTCCACCGTATGACATATTATTAAAGTTGCTAT 1632
Qy 227 AlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGlyAlaGluLeuAlaAspGlu 246
Db 1633 ACAGAGCTGCAAAATTTACATTTGAATTTATTATCATCAAGGTGTACAAATTCGCGGATCAA 1692
Qy 247 TrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTyr 266
Db 1693 TGGAAATGCAGATCAACACCATTCACCAATGTTGAAGTCATCAGTACT-----TAT 1743
Qy 267 TyrLysLeuLysGluAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyrArg 286
Db 1744 TATGACGAGCTATTGGTATATATGAAAGTATATTAAATTTATGACCAAGACATACCAT 1803
Qy 287 ThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTyr 306
Db 1804 AAGAGATTGAATCACCTTAAAGAAATCAGAAAAATCACATGGGATGCTTATAACACATAT 1863
Qy 307 ArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIle 326
Db 1864 CGTCGAGAAATGACCTTAAATGTATTGGATCTTGTGCAACTTTTCCCTTTTATGATATA 1923
Qy 327 LysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThr 346
Db 1924 CGTCGTTT-----CCAGAGAGGTAGAACCTAGAAATNACA 1959
Qy 347 ArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeuArgValGlnPro 366
Db 1960 AGAGAGTTTATACAAAGTTTAGATCATTTAAACAGCACCA----- 2001
Qy 367 AsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeu 386
Db 2002 -----CGGCTATTACTTGGCTG 2019
Qy 387 GluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIleSer 406
Db 2020 TCAGATTATGAGTTATACCGAGAGGTGGCAGAGCGCATTTATTATCAGGTATT--- 2076
Qy 407 AspArgAspAlaProThrTyrSerAsnThr-----IleThrGluThrLeuTyrGly 423
Db 2077 ---CGAGAGCTAAATATATATATCTGGTAACTATTTTACGATGAAAAATATTATGTT 2133
Qy 424 GluArgThrGly---SerProThrThrLysThrIleArgProPheGluSerTyrLysVal 442
Db 2134 AATACAAATAGATTAAGTATAGCAGCTCATTTACATTTATACAGCGCAA-----TATGACT 2190
Qy 443 SerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIleIle 462
Db 2191 CACTTAAGCAATAACCGCTCTTTTCAAACTAGCTGGTATAATAAGTTATACAGTTTA 2250
Qy 463 AsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsnThrLeuLysTyrSerAlaGly 482
Db 2251 ATTCAAAAATCGTATTCACAACTTTTAAAAACGATAAT-----GAATATCAAAAAAT 2304
Qy 483 GlySerLeuSerAsnTyr-----GlnAsnThrThrPhePheGlnPheProArgLys 499
Db 2305 TTTAATGTGAATAATCAAAATGAACCTCAAGAAACTACA----- 2343

Qy 500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHis 519
Db 2344 -----AATATCTCTAATGATTATGGTGTTC-----AACAGCAAAAAATTTCAACAT 2391
Qy 520 IleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnGlnIle 539
Db 2392 AATTATCTCATTTTCCATTAAATCATCCAC-----AAGTTAGAGTTT 2433
Qy 540 LeuAsp-----ThrGlyValLeuGlyTrpThrHisSerSerValAspArgTyr 555
Db 2434 GCTGAGTATTTTCACTCTATATTTGTCATTAGTTGGACACACATAGTTAACTCCCAA 2493
Qy 556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp 575
Db 2494 AATTATATATCAGAAAGTGTGAGTACACAAATCCCATTTGGTAAAGCTTACGAAGTT--- 2550
Qy 576 ThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln 595
Db 2551 ACTAACAAATTCAGTTATAAGAGGACCAAGTTTACAGGTGGAGATTTAATAGAACTTCGT 2610
Qy 596 SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIle 615
Db 2611 GAT-----AAATGTTCTTAAATGTAAAGCT---AGTTCTTTAAAAAATACGCTATA 2661
Qy 616 ArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIle 635
Db 2662 AGTCTATTTTATGCTGCAATTAACGCAATAGCTGTATCAATAGACGTAGGTGATTCGGA 2721
Qy 636 ProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyr 655
Db 2722 GCAGGAGTTCTA-----TTGCAACCTACCTTTTCTAGAAAAAGGAAC 2763
Qy 656 AsnAsn-----LeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSer 670
Db 2764 AATAATTTTACAATTCAGACCTTAATTAAGATTTTCAATATCATACACTTTTAGTT 2823
Qy 671 ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAsp----- 688
Db 2824 GATATTGAATTCGGAAGTGAAGAATTCATATCCATTTGAAGCGAGAGATGATTAT 2883
Qy 689 ValSerAsnSerIleLeuIleAspLysIleGluPheIleProIleThrSerSerMet 708
Db 2884 GAGGAGGAGTGTATCTTTTAAATTTAGATAAATTTAGATTCAAACTATAGATGAAAAAT 2943
Qy 709 HisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePhe 728
Db 2944 -----ACTAATGAATGAATTTAGAGAGCGCAAGAAAGCAGTGAATGATTATT 2994
Qy 729 ThrAsnHisThrLysThrLeu 735
Db 2995 ATAAACGCAACAAACGCTTTG 3015

RESULT 5

US-10-345-020-3

; Sequence 3, Application US/10345020

; Publication No. US20030150018A1

; GENERAL INFORMATION:

; APPLICANT: Jana, Wojciechowska

; APPLICANT: Evgeny, Lewitin

; APPLICANT: Ludmila, Revina

; APPLICANT: Igor, Zalunin

; APPLICANT: Galina, Chestukhina

; TITLE OF INVENTION: NOVEL DELTA-ENDOTOXINS AND NUCLEIC ACID SEQUENCES CODING THEREFOR

; FILE REFERENCE: S-30913B

; CURRENT APPLICATION NUMBER: US/10/345,020

; CURRENT FILING DATE: 2003-01-15

; PRIOR APPLICATION NUMBER: US 60/175,158

; PRIOR FILING DATE: 2000-01-07

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 3

; LENGTH: 4896

```

; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(4896)
; OTHER INFORMATION: Bacillus thuringiensis supsp. finitimus strain VKPM B-1161 (taxon
; OTHER INFORMATION: :29337
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1129)..(4458)
; OTHER INFORMATION: product: Cry28Aa1 delta-endotoxin
; PUBLICATION INFORMATION:
; AUTHORS: Wojciechowska, et al.
; TITLE: Two novel delta-endotoxin gene families cry26 and cry28 from Bacillus
; TITLE: thuringiensis ssp. finitimus
; JOURNAL: FEBS Lett.
; VOLUME: 453
; ISSUE: 12
; PAGES: 46-48
; DATE: 1999-06-18
; US-10-345-020-3

Alignment Scores:
Pred. No.: 1,54e-79 Length: 4896
Score: 849.00 Matches: 229
Percent Similarity: 48.76% Conservative: 106
Best Local Similarity: 33.33% Mismatches: 248
Query Match: 21.94% Indels: 104
DB: 6 Gaps: 25

US-10-783-417-2 (1-735) x US-10-345-020-3 (1-4896)
Qy 68 SerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAla 87
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1210 AGCAGTGATACAGTCGCTGTAGTAAGCGCAGGGATTGTAGTTGGGTACTACTACGACA 1269
Qy 88 GlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIleIle 107
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1270 -----GCCTTTGCATCATTTGTTTAATCCA-----GGTGTGTGTACTTATA 1308
Qy 108 SerPheGlyThrIleIleThrValPheTTPProAlaGlyGluGlnAsp---LysThrVal 126
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1309 TCATTTGGAACCTTGCTCCGCTCCGCTTCTTTGGCCTGATCCAGAGGAAGATCCAAAAAATT 1368
Qy 127 TrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeuThrGluSerIle 146
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1369 TGGTCACAAATTATGAAACAGGAGAGACCTTTAAATCAACAATTTCTACAGCTGTA 1428
Qy 147 LysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGlnSerTyrAsnThr 166
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1429 AAAGAAATAGCATAGCTCATCTAAATGGTTTTAAAGATGTATTAAACGTACTATCAAGA 1488
Qy 167 AlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeuProSerSer 186
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1489 GCATTTAATGATGGAAGAGA-----AATCCAAGTGCA 1521
Qy 187 AlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPhe 206
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1522 -----AATAGTCCAGATTGGTATCATCAGAGATTTGAAACGCTCATTTCAATTTT 1572
Qy 207 IleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeuProIleTyr 226
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1573 GTAAGCAATATGCCACAACTCCAACTTCCCACGTATGACACATTATTATTAAGTTGCTAT 1632
Qy 227 AlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGlu 246
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1633 ACAGAGCTGCAAAATTACATTTGAATTTTATTATCAAGAGGTGTACAAATTCGCGGATCAA 1692
Qy 247 TrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTyr 266
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1693 TGGAAATGCAGATCAACACCAATTCCCAATTTGTTGAAGTCATCAGGTACT-----TAT 1743
Qy 267 TyrLysLeuLysGluAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyrArg 286

```

Db 2662 AGTCTATTTATGCTGCAATAACGCAATAGCTGTATCAATAGACGTAGGTATCCGGA 2721
Qy 636 ProGlyValIleGlyLeProGlnArgLeuAsnThrPheSerGlyThrAsnTyr 655
Db 2722 GCAGGAGTTCTA-----TTGCAACCTACCTTTCTAGAAAGGAAC 2763
Qy 656 AsnAsn-----LeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSer 670
Db 2764 AATAATTTTACAATTCAGACCTTAATACTATAAGGATTTCAATATCATACACTTTTAGTT 2823
Qy 671 ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAsp----- 688
Db 2824 GATATTGAATTAACCCAGAGTGAAGAAATTCATATCCATTTGAAGCAGAGGATGATTAT 2883
Qy 689 ValSerAsnSerIleLeuIleAspLysIleGluPheIleProIleThrSerSerMet 708
Db 2884 GAGGAGGAGTGATCTTTTAAATGATAAATTAGAGTTCAAACTATAGATGAATAATTAT 2943
Qy 709 HisGlnAsnArgGlnLysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePhe 728
Db 2944 -----ACTAATGAATGAATTTAGAGAGGCAAGCAAGAGCAAGTGAATGATTATTT 2994
Qy 729 ThrAsnHisThrLysThrLeu 735
Db 2995 ATAAACGCAACAAACGCTTTG 3015

RESULT 6

US-10-342-821-3
; Sequence 3, Application US/10342821
; Publication No. US20030154510A1
; GENERAL INFORMATION:
; APPLICANT: Jana, Wojciechowska
; APPLICANT: Evgeny, Lewitin
; APPLICANT: Ludmila, Revina
; APPLICANT: Igor, Zalumin
; APPLICANT: Galina, Chestukhina
; TITLE OF INVENTION: NOVEL DELTA-ENDOTOXINS AND NUCLEIC ACID SEQUENCES CODING THEREFOR
; FILE REFERENCE: S-30913C
; CURRENT APPLICATION NUMBER: US/10/342,821
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US 60/175,158
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 3
; LENGTH: 4896
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(4896)
; OTHER INFORMATION: Bacillus thuringiensis supsp. finitimus strain VKPM B-1161 (axon
; OTHER INFORMATION: :29337
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1129)..(4458)
; OTHER INFORMATION: product: Cry28Aa1 delta-endotoxin
; PUBLICATION INFORMATION:
; AUTHORS: Wojciechowska, et al.
; TITLE: Two novel delta-endotoxin gene families cry26 and cry28 from Bacillus
; TITLE: thuringiensis sep. finitimus
; JOURNAL: FEBS Lett.
; VOLUME: 453
; ISSUE: 12
; PAGES: 46-48
; DATE: 1999-06-18
US-10-342-821-3

Alignment Scores:

Pred. No.: 1.54e-79 Length: 4896
Score: 849.00 Matches: 229
Percent Similarity: 48.76% Conservative: 106
Best Local Similarity: 33.33% Mismatches: 248

Query Match: 21.94% Indels: 104
DB: 6 Gaps: 25
US-10-783-417-2 (1-735) x US-10-342-821-3 (1-4896)
Qy 68 SerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAla 87
Db 1210 AGCAGTGATACAGTCGCTGTAGTAAGCGAGGAGTGTGTTGGGTACTATATCTGACA 1269
Qy 88 GlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIleIle 107
Db 1270 -----GCCTTGTCATCATTTGTTAATCCA-----GGTGTGCTACTTATA 1308
Qy 108 SerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGlnAsp---LysThrVal 126
Db 1309 TCATTTGGAACCTTGGCTCCGTTCTTTGGCTGATCCAGAGGAAGATCCAAAAAAT 1368
Qy 127 TrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeuThrGluSerIle 146
Db 1369 TGTGTCAAATTTATGAAACACGAGGAAGACCTTTTAAATCAACAATTTCTACAGCTGTA 1428
Qy 147 LysGlnLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGlnSerTyrAsnThr 166
Db 1429 AAGAATAAGCATTAGCTCATCTAATGTTTAAAGATGATTATTAAGTACTATGAAGA 1488
Qy 167 AlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeuProSerSer 186
Db 1489 GCATTTAATGATTGGAAGACA-----AATCCAAGTGCA 1521
Qy 187 AlaLeuGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPhe 206
Db 1522 -----AATACTGCCAGATTGGTATCACAGAGATTGAAAAACGCTCATTTTCA 1572
Qy 207 IleArgGluIleProGlyPheGlnLeuGluThrLysThrLeuLeuLeuProIleTyr 226
Db 1573 GTAAGCAATATGCCAACACTCCACGATGACACATTTATTATTAAGTTGCTAT 1632
Qy 227 AlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGlu 246
Db 1633 ACAGAAAGTCGCAAAATTTACATTTTGAATTTATACATCAAGGTGTACAATTCGCGGACAA 1692
Qy 247 TrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTyr 266
Db 1693 TGGAAATGCAGATCAACCCACATTCACCAATGTTGAAGTCATCAGGTACT-----TAT 1743
Qy 267 TyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyrArg 286
Db 1744 TATGACGAGCTATTGGTATATATTGAAAGATATATTATTAATTCACCAAGACATACCAT 1803
Qy 287 ThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTyr 306
Db 1804 AAAGGATTGAATCACCTTAAGAAATCAGAAAAATACATCGGATGCTTATAACACATAT 1863
Qy 307 ArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIle 326
Db 1864 CGTCGAGAAATGACCTTAATTTGATGATCTGTGCGCAACTTTCTCTTTTATGATATA 1923
Qy 327 LysArgTyrArgAspSerIleGlyGluValLysGlyIleLysAsnGluLeuThr 346
Db 1924 CGTCGTTTT-----CCAAAGAGAGTAGAATAGATTAATAA 1959
Qy 347 ArgGluIleTyrThrThrGluIleAsnPheAspAspLeuProGlnLeuArgValGlnPro 366
Db 1960 AGAGAGTTTATACAGTTTAGATCATTTAACAGCACCA----- 2001
Qy 367 AsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeu 386
Db 2002 -----GGGCTATTCTTGGCTG 2019
Qy 387 GluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyLeuSer 406
Db 2020 TCAGATATTGATTATACACGAGAGTGTGGCAGAGCGGATTTATTTATCAGGTATT--- 2076


```

QY 407 AsnArgAspAlaProThrTyrSerAsnThr-----lleThrGluThrLeuTyrGly 423
Db 2077 ---CGAGAGCTAAATATATATCTGGTAATCAATTTTTCAGCATGAATAATATTATGGT 2133

QY 424 GluArgThrGly---SerProThrThrLysThrIleArgProPheGluSerTyrLysVal 442
Db 2134 AATACAAATAGATTAAGTAAGCAGCTCATTAATATACCGCGGAA---TTTATGACT 2190

QY 443 SerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIleIle 462
Db 2191 CACTTAAGCATAAACCGCTCTTTTCAACCAATAGCTGGTATAATAGTTATACAGTTTA 2250

QY 463 AsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsnThrLeuLysTyrSerAlaGly 482
Db 2251 ATTCAAAAATCGTATTCACAACTTTTAAAAACGATAAT-----GAATATCAAAAAAT 2304

QY 483 GlySerLeuSerAsnTyr-----GlnAsnThrThrPhePheGlnPheProArgLys 499
Db 2305 TTTTAATGTGAATAATCAAAATGAACCTCAAGAAACTACA----- 2343

QY 500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHis 519
Db 2344 -----AATATCTCTAATATGATTTATGTGTGTTCA---AACAGCCAAAATTCACACAT 2391

QY 520 IleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuIle 539
Db 2392 AATTATATCTCATTTTCCATTAATCATCCAC-----AAGTTAGATTTT 2433

QY 540 LeuAsp-----ThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyr 555
Db 2434 GCTGAGTATTTTCACTCTATATTTGCTATAGTTGGACACACATATGTAAACTCCAA 2493

QY 556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp 575
Db 2494 AATTAAATATCAGAAAGTGTGAGTACAAATCCATTCCTGGTAAAGCTTACGAAGTT-- 2550

QY 576 ThrAsnSerLysValIleGluGlyProGlyHisThrGlyAsnLeuValTyrLeuGln 595
Db 2551 ACTAAACATTCAGTTATTAAGAGGACCAAGGTTTACAGGTGGAGATTTTAATAGAACTTCGT 2610

QY 596 SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIle 615
Db 2611 GAT-----AAATGTTCTATTAATGTAAAGCT---AGTCTTTTAAAAAATACGCTATA 2661

QY 616 ArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIle 635
Db 2662 AGTCTATTTTATGCTGCAATATACGCAATAGCTGTATCAATAGACGTAGGTATCCGA 2721

QY 636 ProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyr 655
Db 2722 GCAGGAGTTCTA-----TTGCAACCTACCTTTTCTAGAAAAGGGAAC 2763

QY 656 AsnAsn-----LeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSer 670
Db 2764 AATAATTTTACAATTCAGACCTTAACATAAGGATTTTCAATATACATACACTTTTATGT 2823

QY 671 ThrValThrProLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAsp 688
Db 2824 GATATTGAATATCCGNAAGTGAAGAATTCATATCCATTGAGCGAGGATGATAT 2883

QY 689 ValSerAsnSerIleLeuIleAspLysIleGluPheIleProIleThrSerSerMet 708
Db 2884 GAGGAGGAGTATCTTTTAAATGATAAATTAGAGTTTCAACCTATAGATGAATAAT 2943

QY 709 HisGlnAsnArgGluLysGlnLysLeuLeuThrIleGlnThrLysIleAsnThrPhePhe 728
Db 2944 -----ACTAATGAATGAATTTTAGAGAAGGCAAAAGACGATGAATGATTTATTT 2994

QY 729 ThrAsnHisThrLysThrLeu 735
Db 2995 ATAAACGCCAACAAACGCTTTG 3015

```

RESULT 7

```

US-10-929-754-2
; Sequence 2, Application US/10929754
; Publication No. US20050124803A1
; GENERAL INFORMATION:
; APPLICANT: ABDULLAH, MODH AMIR
; TITLE OF INVENTION: INSECTICIDAL CRYSTAL PROTEINS WITH ENHANCED TOXICITY
; FILE REFERENCE: 22727/04179
; CURRENT APPLICATION NUMBER: US/10/929,754
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 60/498,826
; FILING DATE: 2003-08-29
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 2
; LENGTH: 3684
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-10-929-754-2

```

```

Alignment Scores:
Pred. No.: 4,93e-70 Length: 3684
Score: 759.00 Matches: 234
Percent Similarity: 47.40% Conservative: 122
Best Local Similarity: 31.16% Mismatches: 267
Indels: 128
Query Match: 19.62%
DB: 9 Gaps: 34

```

US-10-783-417-2 (1-735) x US-10-929-754-2 (1-3684)

```

QY 23 AsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnProLeuGln 42
Db 154 AATATGAATTCAGC-----TATCCGTTAGCGAATGACTTCAAGGGTCAATGAA 204

QY 43 AsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGlyAspAsn 62
Db 205 AACACAGCACTATAAAGATTGCTAGCCATGTGGAATAACCAACAGTATGCGGTTAT 264

QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSer 82
Db 265 -----CCAGCTGCGATTAATCTCTTCAGTTAGTACCCTTTAAAGTAGCT 312

QY 83 GlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
Db 313 GGAGCTATC-----CTTAAATTTGTAAACCCACCTGCAGGTACTGTC 354

QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGly---Glu 121
Db 355 TTAACCGTACTT-----AGCGCGTGTCTCTATTTCTTGGCCGCACTAATCTCCA 405

QY 122 GlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThrPro 141
Db 406 ACGCTGAAGAGTTTGGATGATTTTCATGACCAATACAGGAATCTTATGATCAAACT 465

QY 142 LeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeu 161
Db 466 GTAACAGCTTATGTACGAACAGATGCAAAATGCAAAATGACGGTTGTGAAGATTTT 525

QY 162 GlnSerTyrAsnThrAlaLeuAspAspTyrPargLysLeuLysArgLeuGlnAlaProGly 181
Db 526 GATCAATATACAAATAATTTTAACTTGGAAAGA----- 561

QY 182 LeuProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGluAsn 201
Db 562 GAGCTAATTAACAGTCTCTAGAACAGCAGTAATACT-----CAATTT---AAC 609

QY 202 ValHisAspPheIleArgGluIleProGlyPhe-----GlnLeuGluThrTyrLys 219
Db 610 TTAACAGTGCACAACTTCGAGAGACCGCAGTTTATTTTAGCACTTAGTAGGTTATGAA 669

QY 220 ThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGln 239
Db 670 TTATTGTTATTACCAATATACGCACAAAGTAGCAAAATTTCAATTTTACTTTTAAAGAGAT 729

```


QY	240	GlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluProAsn	259
DB	730	GGCCTCATAAATGCACAAGAATGGTCT-----TTAGCACGTA	768
QY	260	AlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsn	279
DB	769	GCTGGT-----GACCAACTATATAACACTATATGGTCAGTACACTAAAGAATATATTGCA	822
QY	280	TyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLys	299
DB	823	CATAGCATTCATCGTTATATAAAGGTTTAGATGTACTTTAGAAATAAATCTAATAGACAA	882
QY	300	TrpSerIlePheAsnAspTyrArgTyrMetThrIleThrValLeuAspThrIleSer	319
DB	883	TGGATTACGTTTAAATGATTATAAAGAGAGATGACTATTCAAGATTAGATATACTCGCT	942
QY	320	GlnPheSerLeuTyrAspIleLysArgTyr---ArgAspSerIleGlyGlyIleGluVal	338
DB	943	CTTTTTCGCCAGTTATGATCCACGTCGATACCCTCGCGACAAAAATAGATAATACGAAC	1002
QY	339	LysGlyIleLysAsnGluLeuThrArgGluIleTyrThrGluIleAsnPheAspArg	358
DB	1003	TCA-----AAACACAGAAATTACAGAGAGATTATACAGCT-----	1038
QY	359	LeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAla	378
DB	1039	TTAGTAGAATCTCCTCTAGTAAATCTATAGCAGCACTGGAGGCGACACTTACACGA	1095
QY	379	SerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsnPhe	398
DB	1096	GATGTTTCATTATTCATCTGGCTAAAGAGAGTAGATTTCTGGACC---AATACTATATAT	1152
QY	399	GlyAsnArgLeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThr	418
DB	1153	CAAGATTTAAGATTTTATCTGCCAATAAAATGGGTTTCATATACAAATCTCTGCA	1212
QY	419	-----GluThrLeuTyrGlyGluArg---ThrGlySerProThrThrLysThrIle	434
DB	1213	ATCCAGAAAGTGGAAATTATGAAGTTCGTGGTTTGGTTCAATCTTACTCATCAAT	1272
QY	435	ArgProPheGluSer---TyrLysValSerIleValThrAspArgGlnSerProVal	453
DB	1273	CAACTTAAATCTAAATGTTTATAAACTTCTATC---ACAGATACTAGTCTCCCC	1323
QY	454	SerProIleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSer	473
DB	1324	-----TCT	1326
QY	474	AsnAsnThrLeuLys-----TyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsn	490
DB	1327	AATCGAGTTACAAAATGGATTTCTACAAAATGTATGGTACTCTTGCTCTTATATTC	1386
QY	491	-----ThrThrPhePheGlnPheProArgLysLys	500
DB	1387	AATAAACAACCAACTCCTGAAGGTTTAAAGGACCACACATTTTGGATTTTCAACAAATG	1446
QY	501	AspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnIleTyrSerHisIle	520
DB	1447	AAC-----ACACCTAAT---CAACCAACTGTAATGATTATATAGGCATATT	1488
QY	521	LeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeu	540
DB	1489	TTAAGCTAT-----ATAAACTGATGTTATA	1515
QY	541	AspThrGly-----ValLeuGlyTrpThrHisSerSerValAspArgTyrAsn	556
DB	1516	GATTATAACAGTAACAGGGTTTCATTTCCTGGCAGCATATAGATGTTGGACCTTAAT	1575
QY	557	AlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyLysAsnLeuAspThr	576
DB	1576	CAAAATATACAGATGCTATCACACAGATTCGCGCGTAAATCTACTTCTTGAATGCA	1635

```

Qy 577 AsnSerIysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSer 596
      ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1636 ACAGCTAAAGTAATCAAGGACCGCTGCTCATACAGGGGGGGATCTTAGTTGCTCTTCAAGC 1695
Qy 597 -----GlnGlyArgLeuGluIleThrCysGluThrPro-----AsnSer 609
      :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1696 AATGGTACTCTATCAGCGAGAATGGAGATCCAATGTAAACCAAGTAGTATTTTAAATGATCCT 1755
Qy 610 ThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuPro 629
      :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1756 ACAGAGATTACGGATTACGCATACGTTATGCTGCAATAGTCCA-----ATTGTATTG 1809
Qy 630 AsnIleSerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsnAsnThr 649
      :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1810 AATGTATCATATGATTACAAAGGAGTTCTTAGAGGAAACAACGATTAGTACAGAATCTACG 1869
Qy 650 PheSerGlyThrAsnTyr-----AsnAsnLeuGlnTyrClyAspPheGlyTyrPhe 666
      :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1870 TTTTCAAGACCTAATAATATAATPACCTACAGATTTTAAATATGAAAGATTAGATACAA 1929
Qy 667 Gln---PheProSerThrValThrLeuProLeuAsnArgAsn-----IleProPheIle 683
      :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1930 GATCCCTTTTGATGCAATTGTACCGATGAGATTATCTTAATCAACTGATACTATAGCT 1989
Qy 684 PheAsnArgAlaAspVal---SerAsnSerIleLeuIleIleAspIysIleGluPheIle 702
      :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1990 ATTCAACCAATTAAACATGACTTCAAAATATCAAGTGATTATTGACAGAATCGAAATTA 2049
Qy 703 ProIleThrSerSerMetHisGlnAsnArgGluLysGlnIysLeuGluThrIleGlnThr 722
      :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 2050 CCAATCACTCAATCTCTATTAGTAGACACAGAGAACCAAAATTTAGAAATCAGAACGAA 2109
Qy 723 LysIleAsnThrPhePheThrAsnHisThrLys 733
      :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 2110 GTTGTGAATGCACTGTTTACAAATGACGGAAA 2142

RESULT 8
US-10-929-754-4
; Sequence 4, Application US/10929754
; Publication No. US20050124803A1
; GENERAL INFORMATION:
; APPLICANT: ABDULLAH, MODH AMIR
; APPLICANT: DEAN, DONALD H.
; TITLE OF INVENTION: INSECTICIDAL CRYSTAL PROTEINS WITH ENHANCED TOXICITY
; FILE REFERENCE: 2727/04179
; CURRENT APPLICATION NUMBER: US/10/929,754
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 60/498,826
; PRIOR FILING DATE: 2003-08-29
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patent In Ver. 3.2
; SEQ ID NO 4
; LENGTH: 4391
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-10-929-754-4

Alignment Scores:
Pred. No.: 4,48e-64 Length: 4391
Score: 705.00 Matches: 225
Percent Similarity: 43.38% Conservative: 119
Best Local Similarity: 28.37% Mismatches: 278
Query Match: 18.22% Indels: 171
DB: 9 Gaps: 29

US-10-783-417-2 (1-735) x US-10-929-754-4 (1-4391)

Qy 1 MetAsnGlnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
      :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 719 ATGCATTATTATGGGAATAGCATGATATGACATATTA-----ArgTyrProTyr 757
Qy 21 PheProAsnArgAsnSerAsnAspSer-----ArgTyrProTyr 33
      :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

```

758 -----AATGCTTCATCAAAACGATTCAAAACATGTCTAATACTTATCCGAGGTATCCGTTA 811
Db
34 ThrAsnAsnProAsnGlnProLeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCys 53
Qy
812 GCAAAATCCACAACAAGATTAAATCAAAATACAAATATATAAGATGGCTTAATGTATGT 871
Db
54 GlnGlyAsnThrGlnTyrGlyAspAsnPheGluThrPheAlaSerAlaAspThrIleAla 73
Qy
872 GAAGGTAT-----CATATAGAAATCTCTAGAAGACGAGGTAGAGCT 916
Db
74 AlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThr 93
Qy
917 GGACTTGGTAAGGATAGGTATAGTATAGTACAAATCGTAGGGTCTTTGGTGGT---TCT 973
Db
94 SerIleSerGlyProIleGlyIleIleGlyAlaIleIleSerPheGlyThrLeuIle 113
Qy
974 ATTATTAGATCAAAATGGATG-----TTTACCAAAATTTCA 1012
Db
114 ThrValPheTyrProAlaGlyGluGlnAspLysThrValTyrGlnPheIleLysMet 133
Qy
1013 GAGCTACTTGGCCAGAGGATGATACCAGCAATACACTTGGCAAGATATATGAATCAT 1072
Db
134 GlyGluIlePheValAspThrProLeuThrGluSerIleLysGlnLeuLysLeuGlnThr 153
Qy
1073 GTAGAGATCTTATAGACAAACGAATAACTGAGTTATACGAGAAATGCAATTAGAACA 1132
Db
154 LeuGluGlyPheArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLys 173
Qy
1133 TTAGCAGATTACAGGGTAAAGTTGATGATTAACAATGGTTGGAAGAAATGGAAGAC 1192
Db
174 LeuLysArgLeuGlnAlaProGlyLeuProSerSerAlaLeuGlnGlnAlaAlaLeu 193
Qy
1193 -----GATCCAAAATCTACAGTAATTTAAGC----- 1219
Db
194 ThrLeuLysIleArgPheGluAsnValHisAsnAspPhe-----IleArgGluIle 210
Qy
1220 ACCTTAGTACCAGTTTACGGCTCTTGATTGAGATTTTAATGTGCTATNAGGACAGTT 1279
Db
211 -----ProGlyPheGlnLeuGluThrTyrLysThrLeuLeuLeuProIle 225
Qy
1280 AATAATCAGGGAGTCCAGGTTATAGTTA-----CITTTATTCCCTGTC 1324
Db
226 TyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAsp 245
Qy
1325 TATGCAAAATTCGGAATCTGCAATTTACTTTTATTACGGATGCTCAGATTTATGGAGAT 1384
Db
246 GluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAsp 265
Qy
1385 AATGG-----TGGAGCGCAGAGCTAATGCTCGTGATAAT 1420
Db
266 TyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyr 285
Qy
1421 TATTACAAATACAAATTAGAGAAACAAAGGAATATACAGAAATATGTATAAATGGTAT 1480
Db
286 ArgThrGlyLysAsnLeuArgAspGluProAsnMetLysTyrSerIlePheAsnAsp 305
Qy
1481 AATAAGGGTTAAATGATTTTAGAACAGCAGGT-----CAATGGTAAACTTTAATCGT 1534
Db
306 TyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyrAsp 325
Qy
1535 TATCGTAGAAGATGACTCTTACTGTATTAGATATTATCAATGTTCCCTATTATATGAC 1594
Db
326 IleLysArgTyrArgAspSerIleGlyGlyLeuValLysGlyIleLysAsnGluLeu 345
Qy
1595 GCAGATTATCTCAGAA-----GTAAMAAACCGAACTA 1630
Db
346 ThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeuArgValGln 365
Qy
1631 ACTAGGGAATTTATTCAGATGTTTATTAATGGGGAGATATATGCACTTATGACTCCTTAT 1690
Db
366 ProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPheLysLeuPheSerPhe 385
Qy
1691 TTTTCTTTTGAGAAAGCTGAATCACTTTATACAGGGGACCC---CATCTCTTCACTTGG 1747
Db
386 LeuGluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIle 405
Qy
1748 CTAAGAGGATTTGCAATTTGTAAACCAATTCATTTCTTATTTGACATTTTATCAGGTGT 1807
Db
406 SerAsnArgAspAlaProThrTyrSerAsnThrIleThrGlu----- 419
Qy
1808 CAAATAAGTATTTCTTATACTAATAATTTCTAGTATTACGAGGGCTCTTTTAGGGACAG 1867
Db
420 ---ThrLeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGlu 438
Qy
1868 GACACAGATTATGGT-----GGGACTTCTTCTACCATTAATATT---CCATCAAT 1915
Db
439 SerTyrLysValSerIleValThrAspArgGln-----SerProPro 452
Qy
1916 TCGTATGATATAATTTATGACCGGAAATATGATGAATATATTTATCTTGGGGTGATCCT 1975
Db
453 ValSerProIleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySer 472
Qy
1976 GTAATATTACAAAATGAATTTTCTGTAACA-----GATAATAAT 2017
Db
473 SerAsnAsnThrLeuLysTyrSerAlaGlySerLeuSerAsnTyrGlnAsnThrThr 492
Qy
2018 TCTTCAAAAGAATTAATTTATGGTGCACACAGCAATAAACCTGTGTTCTCGACAGAT 2077
Db
493 PhePheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerPro 512
Qy
2078 TTTGATTTTCTCACTAATAAAGAGGGAATCAGTTTAGCA----- 2116
Db
513 AsnPheAsnAsnTyrSerHisIleLeuSerHisPheSerLeu----- 526
Qy
2117 AAATATGAATTAATAATCATATTTATCTATATGTTAAATTAATGGGGAACGTTTGGT 2176
Db
527 -----PheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGly 543
Qy
2177 CAGAAACGTACGTGTTATCGTT----- 2200
Db
544 ValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIle 563
Qy
2201 -----GCTTTTACACATAGTAGTGTGTCCTAATAATACCATTCGACGCAATAAAAT 2254
Db
564 ThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGluGly 583
Qy
2255 ACGCAATTCCTGTAGTAAAGAGCTTCGAGTATAAATGATGATTCATTCATGAAAGGT 2314
Db
584 ProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIleThr 603
Qy
2315 CCCGATTTACGGAGGAGATTTGGTAAAGATGAGAGCAGATTCAGGTTTAACTATGCGT 2374
Db
604 CysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn--- 622
Qy
2375 TTTAAAGCTGAATTTATAGATAAAAAATATCGTGTTCGAATACGTTATAAATGTAACACTAC 2434
Db
623 -----GlyAlaGly-----AsnThrLeuPro 629
Qy
2435 AGTTCTAAATTAATACAGAAAATGGAAGGGAGGTATATACAAACAAATTCAC 2494
Db
630 AsnIleSerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThr 649
Qy
2495 AATATTTCTCCACA----- 2509
Db
650 PheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGln---Phe 668
Qy
2510 -----TATGAGGCTTTTCTTATTAGTCTTTT 2539
Db
669 ProSerThrValThr-----LeuProLeuAsnArgAsnIleProPheIlePheAsn 695
Qy
2540 ACTATAACTACGACAGAAAATATATTTGATTTGACAATGGAGGTAACTACGTATGGT 2599
Db
686 ArgAlaAspValSerAsn---SerIleLeuIleIleAspLysIleGluPheIlePro-Il 704
Qy
2600 AGACAGTTTGTGGAAGATATACCATCTCTTATTATAGATAAAATCGAATTCCTCCCACT 2659
Db


```
Db 529 GATAGTTATTACGCAATACATGCGCATCTTTTCGAGTGACAAATTTTGAAGTACCATTTC 588
Qy 223 LeuProIleTyrAlaGlnAlaAAsnPheHisLeuAsnLeuGlnGlnGlyAlaGlu 242
Db 589 CTTACAGTATATACACAGCGCAGCAACCTTCATTACTGTTATTAAAGACGCTTCAATT 648
Qy 243 LeuAlaAspGluTrpAsnAlaAAsPheHisProSerGlnIleGluProAsnAlaGlyThr 262
Db 649 TTTGGAGAGAAATGGGGATGG-----TCTCAACACCACT 681
Qy 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
Db 682 ATTAATAACTATTATGATCGTCAAACTTATCTGCAGAAATATTCTGATCACTGTGTA 741
Qy 283 AsnThrTyrArgThrGlyLeuLysAsnLeuAspGluProAsnMetLysTrpSerIle 302
Db 742 AAGTGTGTGAACTGTTTGTGCAAAATTAAGGCGCAGCGCTAAACAATGGTGCAC 801
Qy 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db 802 TATAACCAATTCGTTAGAGAAATGACACTGACGGTTTGTAGATGTTGTTCATTATTCCCA 861
Qy 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyLeuValLysGlyIleLys 342
Db 862 AATTATGACACACGACGCTAC-----CCAAATGGAACGAAGCA----- 900
Qy 343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362
Db 901 ---CAACTAACAGGAGAGTATATACA-----GATCCACTGGCGCGGTA 942
Qy 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
Db 943 AAGCTGTCTTCAATTGTTCTCTGATGACAAAGCACCTTCTTTCGAGGTGATAGATCA 1002
Qy 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
Db 1003 TCCGTTATTTCGACCAACC---CATGTATTGTTATATATAACGGGACTCCACAGTGTATACA 1059
Qy 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408
Db 1060 CAATCAAGAGCAATTCCTCCGCTCGCTATATAGACATTGGGCTGTCATCAATAAGC 1119
Qy 409 AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySer 428
Db 1120 TATCATCGGATTTTGTAGTATATATATAAACAGATGTATGGAATAATCAATAATCTA 1179
Qy 429 ProThrThrLysThrIleArgProPheGluSerTyr-----Lys 441
Db 1180 CACAGCACTAGTACCTTT---GATTTTACGAATTATGATATTTTACAAGACGTTATCAAAA 1236
Qy 442 ValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIle 461
Db 1237 GATCGGCTGCTCCTCGATATGTTTTCCTCGTTATACGTTATATA-----TTTTTGGGA 1290
Qy 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
Db 1291 ATGCCAGAAGTCAGTTTTCATGTTGTTATGTTTATGTTTATGTTTATGTTTATGTTTAAAG 1350
Qy 479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr 491
Db 1351 TATAATCCGGTTTCCAAAGATATATATAGCGGG----- 1383
Qy 492 ThrPheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSer 511
Db 1384 -----ACAAGAGATTCGGAATTTAGAAATTCATCCTCCAGAAACTTCA 1422
Qy 512 -----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPhe---SerLeuPheThr 528
Db 1423 GATCAACCAAAATTCAGTCATATACCATAGATTATGTCATATACAAAGTATTCGCCGG 1482
Qy 529 TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTrpThr 548
```

```
Db 1483 ACGGTTCAACTACCGGATTA-----GTACCTGTATTTTCTTGGACA 1524
Qy 549 HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla 568
Db 1525 CATCGGAGTCCGATCTTATAAATGCAGTTTCATTTCAGATAAAATTTACTCAGATTCGGGTC 1584
Qy 569 IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu 582
Db 1585 GTAAGGTTTCTGATTTGGCTCCCTCTATACAGAGGGCCAAATAATACCGTTGTATCG 1644
Qy 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGly----- 598
Db 1645 GGTCCTCGATTTACAGGGGGGGGATATAATAAGTAATAAGAAATAGTAGTAAATATATATCA 1704
Qy 599 -----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616
Db 1705 CATATGCGTGTAAAAATTTCA-----GACATTAAACAAGAATATAGTAGG 1752
Qy 617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
Db 1753 ATTCGGTATGCTTCC-----GCTAATAATACTGAATTTTATATAAAATCCTTCTGAAGAA 1806
Qy 637 GlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn 656
Db 1807 AACGTTAAATCT---CACGCTCAAAAACATATGAATAGAGGTGAAGCTTTTAAACATATAAT 1863
Qy 657 AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValThr 673
Db 1864 AAATTTAATTTATGCGACTTTGCCCCCTATTAAATTTACGACACCGAACCTTTTCATTACT 1923
Qy 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693
Db 1924 CTAGGGCT-----ATATTTCAAGCGGAAGACTTTCTTGGGAATTCAA 1965
Qy 694 LeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGlu 713
Db 1966 GCTTATATAGCGAATCGAATTCGAATTTATCCAGTAGATGAGACATAT-----GAACGC 2016
Qy 714 LysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsn 730
Db 2017 GAACAGATTTAGACGAGCAGGAGAAAGACGAGTGAATGCCTTGTTTACGAAT 2067

RESULT 11
US-10-606-320-3
; Sequence 3, Application US/10606320
; Publication No. US20040091505A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins with Pesticidal
; FILE REFERENCE: 35718/263948
; CURRENT APPLICATION NUMBER: US/10/606,320
; PRIOR FILING DATE: 2003-06-25
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3633
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: CDS
```

```

; LOCATION: (1)...(3633)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Cry12I8-2
US-10-606-320-3
Alignment Scores:
Pred. No.: 5,34e-64 Length: 3633
Score: 703.00 Matches: 219
Percent Similarity: 44.53% Conservative: 127
Best Local Similarity: 28.19% Mismatches: 289
Query Match: 18.17% Indels: 142
DB: 7 Gaps: 32

US-10-783-417-2 (1-735) x US-10-606-320-3 (1-3633)
QY 5 AsnAspAsnAsnGlnTyrGluIleAspSerHisThrSerProTyrPheProAsnArg 24
DB 10 AATAATCAAAATGAATATGAATATATAGATCGACACCTTCT-----ACTTCT 57
QY 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
DB 58 GTATCCAAATGATTCTTAACAGATACCTTTTTCGGAATGAGCCAAACAAATGCGCTACAAAAT 117
QY 44 ThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
DB 118 ATGATTATTAAGATTATTTAAATAATGTGCGGGAAATGCTAGTAATACCTCGGTTC 177
QY 63 PheGlnThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSer 82
DB 178 CCTGAAGTACTTGTAGCGGACAGAT-----GCAGTAAAGCGCGCAATGATATAGTA 231
QY 83 GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
DB 232 GGTAAATTACTATCAGGTTTAGG-----GTCCCATTTGTT 267
QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGlyGluGln 122
DB 268 GGCCCGATAGTAGTCTTTATATCTCAACTTATTGATATTCTGTGGCCTTCAGGGCAA--- 324
QY 123 AspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu 142
DB 325 ---AAGAGTCAATGGGAGATTTTATGAAACAAGTAGAAGAACTCAATAAACAATA 381
QY 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
DB 382 GCAGATATCCAGGAATAAGCGCTTTCGGAATTAGAGGATTTAGTAATAATACCAA 441
QY 163 SerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
DB 442 TTATATCTAACTGCGCTTGAAGAATGAAAGAA----- 474
QY 183 ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal 202
DB 475 AATCCAAATGGTTCA-----AGAGCCTTACGAGATGTGCGAAATCGAATCGAATCCTG 528
QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
DB 529 GATAGTTATTTAGCGAATACAGCCATCTTTTCGAGTGACAAATTTTGAAGTACCAATTC 588
QY 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGlyAlaGlu 242
DB 589 CTTACAGTATATACACAGCGAGCAACCTTCATTACTGTATTAAAGAGCGCTTCAATT 648
QY 243 LeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
DB 649 TTTGGAGAGCAATGGGATGG-----TCTCAACACCACT 681
QY 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
DB 682 ATTAATAAATATTATGATCGTCAAAATGAAACTTACTCGCAAAATATTCTGATCAGTGTGA 741

```

```

QY 283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTyrSerIle 302
DB 742 AAGTCGTATGAACACTGGTTTACAAAATTAAGAGCAGCGCTTAAACAATGGGTCGAC 801
QY 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
DB 802 TATAACCAATTCCTAGAGAAATGACATGACGGTTTATAGATGTTGTTGCTATTATTTCCCA 861
QY 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIleLys 342
DB 862 AATTATGACACACGCGCTAC-----CCAATGAAACGAAAGCA----- 900
QY 343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362
DB 901 ---CAACTAACAGGAAAGTATATACA-----GATCCACTGGGCGCGGTA 942
QY 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
DB 943 AACGTGCTTCAANTGGTTCCTGGTATGACAAAGCACCTTCTTTCGGAGTGATAGATCA 1002
QY 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
DB 1003 TCCGTTATTCGACCACCC---CATGTATTGATTATATAACGGGACTCACAGTGATACA 1059
QY 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408
DB 1060 CAATCAAGAACCAATTTCTCCGCTCTATATAAGACATTTGGGCTGTCATCAATAAAGC 1119
QY 409 AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySer 428
DB 1120 TATCATCGGATTTTAGTAAATATATATAACAGATGTATGGAACATAAATACTTA 1179
QY 429 ProThrThrThrIleArgProPheGluSerTyr-----Lys 441
DB 1180 CACAGCACTAGTACTCTT---GATTTTACGAATTATGATATTTACAGAGCTTATCAAAA 1236
QY 442 ValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIle 461
DB 1237 GATCGGTCTCTCTGATATTGTTTCTCTGTTATACGTATATA-----TTTTTGA 1290
QY 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
DB 1291 ATGCCAGAAAGTCGAGTTTTTCATGTGTAACCAATTTGAATAATACCAGAAAGACGTAAAG 1350
QY 479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr 491
DB 1351 TATAATCCGGTTTCCAAAGATATTATACGGGG----- 1383
QY 492 ThrPhePheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSer 511
DB 1384 -----ACAAGAGATTCGGAAATTAGAAATTACCTCCAGAAACTTCA 1422
QY 512 -----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPhe---SerLeuPheThr 528
DB 1423 GATCAACCAAAATTAATGATCATATAGCCATAGATTATGTCATATACAAAGTATCCCGCG 1482
QY 529 TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThr 548
DB 1483 ACGGGTTCAACTACCGGAATTA-----GTACCTGTATTTCTTGGACA 1524
QY 549 HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleIleThrMetIleProAla 568
DB 1525 CATCGGAGTCCGATCTTATAAATGCAGTTCATTACAGATAAAATTTACTCAGATTCCGGTC 1584
QY 569 IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu 582
DB 1585 GTAAAGGTTTCTGATTTGGCTCCCTCTATAACAGAGGCGCAAAATAAATACCTGTATCG 1644
QY 583 GlyProGlyHisThrGlyLysAsnLeuValTyrLeuGlnSerGlnGly----- 598
DB 1645 GGTCTCGATTTTACAGGGGGGGGATATAAAGTAATAAGAAATAGAGTAATATATATCA 1704
QY 599 -----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616

```



```

Db 1705 CATATGGTGTAAATTTCA-----GACATTACAAAGAAATATAGTATGAGG 1752
Qy 617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
Db 1753 ATTTCGTATGCTCC-----GCTAATAACTGAAATTTTATATAAATCTCTCTGAAGAA 1806
Qy 637 GlyValIleGlyLeuProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn 656
Db 1807 AACGTTAAATCT---CACGCTCAAAAACATATGAATAGAGTGAAGCTTTAAACATATAAT 1863
Qy 657 AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValThr 673
Db 1864 AAATTTAATATCGAGCTTTGCCCTTATTAATAATTACGACCAACCGAACCCTTTCATTACT 1923
Qy 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693
Db 1924 CTAGGGGCT-----ATATTTGAACCGGAAGACTTTCTTGGNAATTGAA 1965
Qy 694 LeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGlu 713
Db 1966 GCATTATAGACCGAATCGAATTTATCCAGTAGATGAGACATAT-----GAAAGCG 2016
Qy 714 LysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsn 730
Db 2017 GAACAAGATTTAGAACGACGCAAGAAAGCAGTGAATGCCTTGTGTTACGAAT 2067

```

RESULT 12

```

US-10-746-914-3
; Sequence 3, Application US/10746914
; Publication No. US20040210963A1
; GENERAL INFORMATION:
; APPLICANT: Albert L. Lu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE REFERENCE: Activity
; CURRENT APPLICATION NUMBER: US/10746,914
; CURRENT FILING DATE: 2003-12-24
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3633
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3633)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Cry12I8-2
US-10-746-914-3

```

Alignment Scores:

```

Pred. No.: 5,34e-64 Length: 3633
Score: 703.00 Matches: 219
Percent Similarity: 44.53% Conservative: 127
Best Local Similarity: 28.19% Mismatches: 289
Query Match: 18.17% Indels: 142
DB: 8 Gaps: 32

```

US-10-783-417-2 (1-735) x US-10-746-914-3 (1-3633)

```

Qy 5 AsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyrPheProAsnArg 24
Db 10 AATAATCAAAATGAATATGAATATATAGATGCGACACCTTCT-----ACTTCT 57

```

```

Qy 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
Db 58 GTATCCAAATGATTTAAACAGATACCTTTTCGGAATGAGCAACAAATCGCGTACAAAAT 117
Qy 44 ThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
Db 118 ATGATTTATAAGATTAATTTAAATAATGCTCGGGGAATGCTAGTAATACCTCGGTGTA 177
Qy 63 PheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSer 82
Db 178 CCTGAAGTACTTGTAGCGGACAGAT-----CGAGCTAAGCGCGCAATTGATATAGTA 231
Qy 83 GlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
Db 232 GGTAAATTAATCTATCAGGTTTAGG-----GTCCCATTTGTT 267
Qy 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGlyGluGln 122
Db 268 GGGCCGATAGTAGTCTTTTATCTCAACTTATGATATTTCTGTGGCCTTCAGGGCAA--- 324
Qy 123 AspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu 142
Db 325 ---AAGAGTCAATGGGAGATTTTATGGAACAAGTAGAGAACTCATAAATCAAAAATA 381
Qy 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
Db 382 CGAATATGCAAGCAATAAAGCGCTTTCGGAATTAGAAGGATTAGGTAAATAATACCAA 441
Qy 163 SerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
Db 442 TTATATCTAACTGCGCTTGAAGATGGAAGAA----- 474
Qy 183 ProProSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal 202
Db 475 AATCCAAATGGTTCA-----AGAGCCTTACGAGATGCGGAATCGATTTGAAATCCTG 528
Qy 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluTyrThrLysThrLeuLeu 222
Db 529 GATAGTTTATTAGCGCAATACATGCCATCTTTTCGAGTGACAAATTTTGAAGTACCATTC 588
Qy 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGlu 242
Db 589 CTTACAGTATATACACAGCGACCAACCTTCATTACTGTTATTAAAGAGACGCTTCAATT 648
Qy 243 LeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
Db 649 TTTGGAGAAGAATGGGGATGG-----TCTACAACTACT 681
Qy 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
Db 682 ATTAATAACTATTATGATCGTCAATGAAACTTACTGCAGAAATATTCTGATCACCTGTGTA 741
Qy 283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle 302
Db 742 AAGTGGTATGAAACTCGGTTTAGCAAAATTAAGAGCGACGAGCGGTAAACAATGGGTGCGAC 801
Qy 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db 802 TATAACCAATTCGCTAGAGAAATGACACTGACGGTTTGTAGATGTTGTTGCAATTATCCCA 861
Qy 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIleLys 342
Db 862 AATTATGACACACCGCAGTAC-----CCAATGGAACCAAGAACCA----- 900
Qy 343 AsnGluLeuThrArgGluIleTyrThrGluIleAsnPheAspArgLeuProGlnLeu 362
Db 901 ---CAACTAAACAGGGAAGTATATACA-----GATCCACTGCGGCGGTA 942
Qy 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
Db 943 AACGTGTCTTCAATTTGGTTCTGTTATGACAAAGCACCTTCTTTCGGAGTAGTGAAGATCA 1002
Qy 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnPheIlePheTyrThr 393

```



```
Db 1003 TCCGTTATTCGACCACCC---CATGTTATTTGATTATATAACGGGACTCACAGTGTATACA 1059
Qy 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408
Db 1060 CAATCAAGAAGCAATTCCTCCGCTCGCTATATAAGACATTGGCTGGTCATCAATAAAGC 1119
Qy 409 AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySer 428
Db 1120 TATCATCGGATTTTATGATAATATATAAACAAGATGTATGGAACATAATCAAAATCTA 1179
Qy 429 ProThrThrLysThrIleArgProPheGluSerTyr-----Lys 441
Db 1180 CACAGCACTAGTACCTTT---GATTTTACGAATTATGATATTTACAGACGTTATCAAAA 1236
Qy 442 ValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIle 461
Db 1237 GATCGGTGTCCTTGTATTTGTTTCTCTGTTATATAGTATATA-----TTTTTGGG 1290
Qy 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
Db 1291 ATGCCAAGAGTCGAGTTTTCATCGTAAACCAATTGAATAATACAGAAAGACGTTAAAG 1350
Qy 479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr 491
Db 1351 TATAATCCGGTTTCCAAAGATATATAGCGGG----- 1383
Qy 492 ThrPhePheGlnPheProArgLysAspCysAsnLeuValIleAspProGlyCysSer 511
Db 1384 -----ACAAGAGATTCGGAATTAGAAATTACCTCCAGAACTCA 1422
Qy 512 -----ProAsnPheAsnTyrSerHisIleLeuSerHisPhe-----SerLeuPheThr 528
Db 1423 GATCAACCAAAATATGATCATATAGCCATAGATTTATGTCATATACAAAGTATTTCCCGC 1482
Qy 529 TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThr 548
Db 1483 ACGGTTCAACTACCGGATTA-----GTACCTGATTTTCTTGACA 1524
Qy 549 HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla 568
Db 1525 CATCGGAGTCCGATCTTATAAATGCAGTTCATTCAGATAAAATATCATGATTCGCGTC 1584
Qy 569 IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu 582
Db 1585 GTAAAGGTTTCTGATTTGGCTCCCTCTATAACAGAGGGCCAAATAATACCGTTGTATCG 1644
Qy 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGly----- 598
Db 1645 GGTCTCGATTTACAGGGGGGGGATAATAAAGTAATAAGAAATGGAGTAATATATATCA 1704
Qy 599 -----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616
Db 1705 CATATGCGTGTAAANTTCA-----GACATTAACAAGAAATATAGTATGAGG 1752
Qy 617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
Db 1753 ATTGCGTATGCTTCC-----GCTAATAATACTGAATTTTATATAAATCCTTCTGAAGA 1806
Qy 637 GlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn 656
Db 1807 AACGTTAAATCT---CACGCTCAAAAAACTATGAATAGAGGTGAAGCTTTAAACATATAAT 1863
Qy 657 AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValThr 673
Db 1864 AAATTTAAATATGCGACTTTGGCCCTTATTAAATTTACGACCAACCCACCTTCTTACT 1923
Qy 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693
Db 1924 CTAGGGCT-----ATATTGAAGCGGGAAGACTTTCTTGGAAATTGAA 1965
Qy 694 LeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGlu 713
Db 1965 -----ATTTTGAAGCGGGAAGACTTTCTTGGAAATTGAA 1965
```

```
Db 1966 GCTTATATAGACCGAATCGAATTTTATCCAGTAGATGAGACATAT-----GAAGCG 2016
Qy 714 LysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsn 730
Db 2017 GAACAAGATTTTAGAAGCAGCGAAGAAAGCAGTGAATGCTTGTGTACGAAT 2067

RESULT 13
US-10-032-717-28
; Sequence 28, Application US/10032717
; Publication No. US20020151709A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Nicholas B. Duck
; APPLICANT: Xiang Peng
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Theodore W. Kahn
; APPLICANT: Lynn E. Sims
; TITLE OF INVENTION: Genes Encoding No. US20020151709A1el Proteins With
; FILE REFERENCE: 35718/237005
; CURRENT APPLICATION NUMBER: US/10/032,717
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 6613
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Genomic Cry1218-2
US-10-032-717-28

Alignment Scores:
Pred. NO.: 1,47e-63 Length: 6613
Score: 703.00 Matches: 219
Percent Similarity: 44.53% Conservative: 127
Best Local Similarity: 28.19% Mismatches: 289
Query Match: 18.17% Indels: 142
DB: 32

US-10-783-417-2 (1-735) x US-10-032-717-28 (1-6613)
Qy 5 AsnAspAsnAsnGluTyrGluIleAspSerHisThrSerProTyrPheProAsnArg 24
Db 1263 AATAATCAAAATGAATATGAATAATATATAGTACGACACCTTCT-----ACTTCT 1310
Qy 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
Db 1311 GTATCCCAATGATTTCAACAGATACCTTTTCGGAATGAGCCAAACAATGCGGTACAAAT 1370
Qy 44 ThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
Db 1371 ATGGATTAAAGATTTATTTAAATAATGTCGCGGAATGCTAGTGAATACCTGCTTCA 1430
Qy 63 PheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIleValSer 82
Db 1431 CCTGAAGTACTTGTGTAGCGGCAAGAT-----GCAGCTAAGGCGGCAATTGATATAGTA 1484
Qy 83 GlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
Db 1485 GGTAAATTTACTATCAGGTTTAGGG-----GTCCCATTTGTT 1520
Qy 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGlyGluGln 122
Db 1521 GGGCGCATAGTACCTTTTACTCAACTTATGATATCTGCTGCTTACGGGCAA--- 1577
Qy 123 AspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu 142
Db 1578 ---AAGAGTCAATCGGAGATTTTATGGAACAAGTAGAAGAACTCATATAATCAAAAAATA 1634
```

QY 143 ThrGluSerIleLysGlnLeuLysLysGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
Db 1635 GCAGATATGCAAGGAATAAAGCGCTTCGGAATTAGAAGGATTAGGTAATAATTACCAA 1694
QY 163 SerTyrAsnThrAlaLeuAspTyrArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
Db 1695 TTATATCTAACTCGCTTGNAGATGGAAGAA----- 1727
QY 183 ProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGluAsnVal 202
Db 1728 AATCCAAATGGTTCA-----AGAGCCTTACGAGATGCGGAATCGATTTGAAATCCTG 1781
QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
Db 1782 GATAGTTTATTTACGCAATACATCGCATCTTTTCGAGTGACAAATTTTGAAGTACCATT 1841
QY 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGlyAlaGlu 242
Db 1842 CTTACAGTATATACAGGCGACCAACCTTCATTTTACTGTTTAAAGGACGCTTCAATT 1901
QY 243 LeuAlaAspGluTyrPheAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
Db 1902 TTTGGAGAGATGGGATG-----TCTACACCACT 1934
QY 263 SerAspAspTyrTyrLysLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
Db 1935 ATTAATAACTATTATGATCGTCAATGAACTTACTGCGAATATTCTGATCACTGTGTA 1994
QY 283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTyrSerIle 302
Db 1995 AAGTGTGTATGAATCGTTTGTAGCAAAATTTAAAGCGCACGCGCTAAACAAATGGGTGAC 2054
QY 303 PheAsnAspTyrArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db 2055 TATAACCAATTCCTGAGAGAAATGACACTACCGTTTGTAGTGTGTCATTATCCCA 2114
QY 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIleLys 342
Db 2115 AATTATGACACGCGATAC-----CCAAATGGAACGAAGCA----- 2153
QY 343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362
Db 2154 ---CAACTAACAGGGAAGTATATACA-----GATCCACTGGGCGCGTA 2195
QY 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
Db 2196 AAGGTCTTCAATTCCTGCTGCTATATGAAGACATTGGCTGCTCATCAATAAGC 2255
QY 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
Db 2256 TCCGTTATTCGACCAACC---CATGTATTGTATATATAACGGGACTCACAGTGTATACA 2312
QY 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408
Db 2313 CAATGAAGAAGCATTTCTTCGCTGCTATATGAAGACATTGGCTGCTCATCAATAAGC 2372
QY 409 AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySer 428
Db 2373 TATCATCGGATTTTATGATGATAATATATATAAACAAGATGTATGGAACATAATCAAAATCTA 2432
QY 429 ProThrLysThrIleArgProPheGluSerTyr-----Lys 441
Db 2433 CACAGCACTAGTACCTTT---GATTTTACGAATATGATATTTACAGAGCGTTATCAAAA 2489
QY 442 ValSerIleValThrAspArgGlnSerProValSerProIleGlnProHisPheIle 461
Db 2490 GATCGCGTCTCTGTATATTGTTTCTGTTTATATACGATATA-----TTTTTTGGA 2543
QY 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
Db 2544 ATGCCAAGATGCGAGTTTTCATGGTAAACCAATTTGAATAATAACGAAAGACGTTAAAG 2603

QY 479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr 491
Db 2604 TATAATCCGGTTTCCAAAGATATTATAGCGGG----- 2636
QY 492 ThrPhePheGlnPheProArgLysLysAspCysAsnLeuValIleAaspProGlyCysSer 511
Db 2637 -----ACNAGAGATTCGGAATTAGNAATTACCTCCAGAAACTTCA 2675
QY 512 -----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPhe---SerLeuPheThr 528
Db 2676 GATCAACCAAAATATTAGTTCATATAGCATAGATTATGTATATATACACAGTATTTCCCGCG 2735
QY 529 TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThr 548
Db 2736 ACGGGTTCAACTACCGGATTA-----GPACCTGTATTTTCTTGGACA 2777
QY 549 HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMerIleProAla 568
Db 2778 CATCGGAGTCCCGATCTTATAAATGCGAGTTCAATTCAGATAAAATTACTCAGATTCGCGTC 2837
QY 569 IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu 582
Db 2838 GTAAAGGTTTCTGATTTGGCTCCCTCTATAACAGGAGGCCAATAATACCGTTGTATCG 2897
QY 583 GlyProGlyHisThrGlyAsnLeuValTyrLeuGlnSerGlnGly----- 598
Db 2898 GGTCTCGATTTTACAGGGGGGGGATAATAAAGTAATAAGAAATGAGTAGTAATTATATCA 2957
QY 599 -----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616
Db 2958 CATATGCGTGTATAAATTTCA-----GACATTAAACAAGAATAATAGTATAGG 3005
QY 617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
Db 3006 ATTCCGTATGCTTCC-----GCTAATAATACTGAATTTTATATAAATCTCTCTGAAGAA 3059
QY 637 GlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn 656
Db 3060 AAGCTTAAATCT---CACGCTCAAAAAATATGAATAGAGGTGAAGCTTTTAAACATATAAT 3116
QY 657 AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValThr 673
Db 3117 AAATTTAATTTATCGCACTTTGCCCTTATTAATAATTTAGCAACCGAACCTTTTCACTACT 3176
QY 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693
Db 3177 CTAGGGCT-----ATATTTGAACGGGAGACTTCTTCTTGGAAATTGAA 3218
QY 694 LeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGlu 713
Db 3219 GCTTATATACCGCAATCGAATTTATCCAGTAGATGAGACATAT-----GAAGCG 3269
QY 714 LysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsn 730
Db 3270 GAACAAGATTTAGAAAGCAGCGAAGAAAGACGTGAATGCTTGTGTTTACGAAT 3320

RESULT 14
US-10-414-637-28
; Sequence 28, Application US/10414637
; Publication No. US20030177528A1
; GENERAL INFORMATION:
; APPLICANT: Nicholas B. Duck
; APPLICANT: Xiang Peng
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Theodore W. Kahn
; APPLICANT: Lynn E. Sims
; TITLE OF INVENTION: Genes Encoding No. US20030177528A1el Proteins With
; PESTICIDAL ACTIVITY AGAINST COLEOPTERANS
; FILE REFERENCE: 35718/237005
; CURRENT APPLICATION NUMBER: US/10/414.637
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: US/10/032.717

```
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 6613
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Genomic Cyl1218-2
US-10-414-637-28

Alignment Scores:
Pred. No.: 1,47e-63 Length: 6613
Score: 703.00 Matches: 219
Percent Similarity: 44.53% Conservative: 127
Best Local Similarity: 28.19% Mismatches: 289
Query Match: 18.17% Indels: 142
DB: 32

US-10-783-417-2 (1-735) x US-10-414-637-28 (1-6613)

QY 5 AsnAspAsnGluTyrGluIleAspSerHisThrSerProTyrPheProAsnArg 24
D 1263 AATAATCAAAATGAATATGAATATATAGATCGACACCTTCT-----ACTTCT 1310
QY 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGluAsn 43
D 1311 GTATCCAAATGATCTCAACAGATACCCCTTTTGGCAATGAGCCCAACAAATGCGCTACAAAT 1370
QY 44 ThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
D 1371 ATCGATTATAAGATATTTTAAATCTCTCGCGGAATGCTAGTGAATACCTCGGTTC 1430
QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSer 82
D 1431 CCTGAAGTACTGTAGCGGACAAAGAT-----GCAGCTAAGCGCCCAATGATATAGTA 1484
QY 83 GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIle 102
D 1485 GGTAAATTTACTACAGGTTTAGG-----GTCCCAATTTGTT 1520
QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGlyGluGln 122
D 1521 GGCCCGATAGTAGCTTTTACTCAACTTATTGATATTCTGTGGCCCTTCAGGCGCA-- 1577
QY 123 AspLysThrValTyrThrGlnPheIleLysMetGlyGlyIlePheValAspThrProLeu 142
D 1578 ---AAGAGTCAATGGGAGATTTTATGGAACAGTAGAAGACTCAATAATCAAAAATA 1634
QY 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
D 1635 GCAGATATCAAGGAATAAAGCGCTTCGGAATTAGAGATTAGGTAATAATTACCAA 1694
QY 163 SerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
D 1695 TTATATCTAACTCGCTTGAAGATGGAAAGAA----- 1727
QY 183 ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal 202
D 1728 AATCCAAATGGTTCA-----AGAGCCTTAGAGATGTGGAATCGAATCGATTGAAATCCGT 1781
QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
D 1782 GATAGTTATTACGCAATACATGCCATCTTTTCGAGTGCACAAATTTTGAAGTACCATTC 1841
QY 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlyAlaGlu 242
D 1842 CTTACAGTATATACACAGCGGACCAACCTTCATTACTGTTATTAAAGGAGCGCTTCAATT 1901

QY 243 LeuAlaAspGluTyrProAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
D 1902 TTTGGAGAAGAATGGGGATGG-----TCTACAAACCACT 1934
QY 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
D 1935 ATTAATAACTATTATGATCGTCAATGAACCTTACTGAGAAATATCTGATCAGTCTGTGTA 1994
QY 283 AsnThrTyrArgThrGlyLeuLysAsnLeuAspGluProAsnMetLysTrpSerIle 302
D 1995 AAGTGGTATGAACCTGTTTAGCAAAATTAAGAGCGCAGCGCTAAACAAATGGGTCGAC 2054
QY 303 PheAsnAspTyrArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
D 2055 TATAACCAATTTCCGTAGAGAAATGACATGACGCTTTTAGATGTTTGTGCTATTATCCCA 2114
QY 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIleLys 342
D 2115 AATTATGACACAGCGCATC-----CCAATGGAACAGAAAGCA----- 2153
QY 343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362
D 2154 --CAACTAACAGGAAGTATATACA-----GATCCACTGGCGCGGTA 2195
QY 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
D 2196 AACCTGCTCTCAATTTGGTTCTGCTATGACAAAGCACCTTCTTTGCGAGTATAGATCA 2255
QY 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
D 2256 TCGGTTATTCGACCAACC-----CATGATTGATTATATAACGGGACTCACAGTGATATACA 2312
QY 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408
D 2313 CAATCAAGACGATTTCTTCGCTCGCTATATAAGACATTGGCGTGGTTCATCAATAAGC 2372
QY 409 AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySer 428
D 2373 TATCATCGGATTTTAGTGAATAATATAAAACAGATGATGGAACCTAATAATCAAAATCTA 2432
QY 429 ProThrThrLysThrIleArgProPheGluSerTyr-----Lys 441
D 2433 CACAGCACTAGTACCTTT---GATTTCAGAAATTATGATATTTTACAAGACGTTATCAAAA 2489
QY 442 ValSerIleValThrAspArgGlnSerProValSerProIleGlnProHisPheIle 461
D 2490 GATCGGTGCTCTGTGATTTGTTTTCTGGTTATACGTATATA-----TTTTTTGA 2543
QY 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
D 2544 ATGCCAGAGTCGAGTTTTCATGGTAAACCAATTTGAATATATACCAGAAAGACGTTAAAG 2603
QY 479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr 491
D 2604 TATAATCCGGTTTCCAAAGATATTATAGCGGG----- 2636
QY 492 ThrPhePheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSer 511
D 2637 -----ACAAGATTCGGAATTAGAAATACCTCCAGAAACTTCA 2675
QY 512 -----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPhe---SerLeuPheThr 528
D 2676 GATCAACCAATATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCTCGCG 2735
QY 529 TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTrpThr 548
D 2736 ACGGTTTCAACTACCGGATTA-----GTACCTGTATTTCTTTGGACA 2777
QY 549 HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla 568
D 2778 CATCGGATGCGGATCTTTATAAATGCAGTTCTATTTCAGATAAAATTAATCAGATTCGCGTC 2837
QY 569 IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu 582
```



```
Db 2115 AATTATGACACACGACGTAC-----CCAAATGGAACGAAAGCA----- 2153
Qy 343 AsnGluLeuThrArgGluLeuThrThrThrGluLeuPheAspArgLeuProGlnLeu 362
Db 2154 ---CAACTAAACAGGAAGTATATACA-----GATCCACGGCGCGGTA 2195
Qy 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
Db 2196 AAGTGTCCTCAATTGGTCTCGTGTATGACAAAGCACCTCTTCGGAGTGATAGATCA 2255
Qy 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
Db 2256 TCGGTATTTCGACCACCC---CATGTATTGATTATATAACGGGACTCACAGTGTATACA 2312
Qy 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408
Db 2313 CAATCAAGACGATTTCTTCGGCTCGCTATATAAGACATTGGGCTGGTCAATCAATAAGC 2372
Qy 409 AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySer 428
Db 2373 TATCATCGGATTTTGTAGTATAATTATATAAACAAGATGTATGGAACATAATCAAAATCTA 2432
Qy 429 ProThrThrLysThrIleArgProPheGluSerTyr-----Lys 441
Db 2433 CACAGCACTAGTACCTTT---GATTTACGAAATTATGATATTTACAAGACGTTATCAAAA 2489
Qy 442 ValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIle 461
Db 2490 GATCGGTGTCTCTGTATATTGTTTCTCGTTATACGTATATA-----TTTTTGGGA 2543
Qy 462 IleAsnGlnIleGluLeuLysGlySerSerAsnAsn-----ThrLeuLys 478
Db 2544 ATCCACAGAGTCGAGTTTTTCATCGTAAACCAAAATTTGAATAATACCAGAAAGACGTTAAAG 2603
Qy 479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr 491
Db 2604 TATAATCCGGTTTCCAAAGATATATATAGCGGG----- 2636
Qy 492 ThrPhePheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSer 511
Db 2637 -----ACAAGAGATTCGGAATTAGAAATTACCTCCAGAAACTTCA 2675
Qy 512 -----ProAsnPheAsnTyrSerHisIleLeuSerHisPhe-----SerLeuPheThr 528
Db 2676 GATCAACCAAAATTATGAGTCATATAGCCATAGATTATGTCATATACAAAGTATTCGCGCG 2735
Qy 529 TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThr 548
Db 2736 ACGGGTCAACTACCGGATTA-----GTACCTGTATTTTCTTGGACA 2777
Qy 549 HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla 568
Db 2778 CATCGGAGTGCCGATCTTATAAATGCAGTTCATTCAGATAAAATTAATCAGATTCGCGTC 2837
Qy 569 IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu 582
Db 2838 GTAAGGTTCTCATTTGGTCCCTCTAACAAGAGCGGCCAAATAATACCGTTGTATCG 2897
Qy 583 GlyProGlyHisThrGlyLysLeuValTyrLeuGlnSerGlnGly----- 598
Db 2898 GGTCTCGATTTACAGGGGGGGGATATAAAGTAATAAGAAATGGAGTAATATATATCA 2957
Qy 599 -----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616
Db 2958 CATATCGGTGTATAAAATTCA-----GACATTAACAAAGAAATATAGTAGG 3005
Qy 617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
Db 3006 ATTCGGTATGCTTCC-----GCTAATAATACTGAATTTTATATAATCTCTCTGAAGAA 3059
Qy 637 GlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn 656
Db 3060 AACGTTAAATCT---CACGCTCAAAAAAATATGAATAGAGGTGAAGCTTTTAACATATAAT 3116
```

```
Qy 657 AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValThr 673
Db 3117 AAATTTAATTATGCGACTTTGCCCTTATTAATAATTACGACAAACCGAACCTTTTCATTACT 3176
Qy 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693
Db 3177 CTAGGGGCT-----ATATTGAAGCGGAAGACTTTTCTTGGAAATTGAA 3218
Qy 694 LeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGlu 713
Db 3219 GCTTATATAGACCGAATCGAATTTATCCAGTATGATGAGACATAT-----GAAGCG 3269
Qy 714 LysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsn 730
Db 3270 GAACAAGATTTAGAAAGCAGCGAAGAAAGCAGTGAATGCCTTGTGTACGAAT 3320
```

Search completed: December 11, 2005, 23:17:11
Job time : 1528 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 11, 2005, 20:06:37 ; Search time 173 Seconds

(without alignments)
1588.342 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869

Sequence: 1 MNQNDNNEYIIDSHTSPY.....KLETIQTKINTFFNHTKYL 735

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3392430 seqs, 186927314 residues

Total number of hits satisfying chosen parameters: 6784860

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-Q=/cgn2_1/USPTO.spool/US10783417/runat_01122005_141955_24269/app_query.fasta_1.903
-DB=Published Applications NA.New -QWMT=fastap -SUFFIX=rnpbn -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10783417 @CGN 1 139 @runat 01122005_141955_24269
-NCPUS=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOF=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
- 9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
- 10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	703	18.2	3633	US-11-058-727-3	Sequence 3, Appli
2	703	18.2	3633	US-11-108-389-3	Sequence 3, Appli
3	703	18.2	6613	US-11-058-727-18	Sequence 18, Appli
4	703	18.2	6613	US-11-108-389-18	Sequence 18, Appli
5	689.5	17.8	4188	US-11-091-643-5	Sequence 5, Appli
6	678.5	17.5	2022	US-11-058-727-13	Sequence 13, Appli
7	678.5	17.5	2022	US-11-108-389-13	Sequence 13, Appli
8	675	17.4	3621	US-11-058-727-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-11-058-727-3
; Sequence 3, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3633
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:

Sequence 1, Appli
Sequence 17, Appli
Sequence 3, Appli
Sequence 73, Appli
Sequence 79, Appli
Sequence 73, Appli
Sequence 79, Appli
Sequence 79, Appli
Sequence 81, Appli
Sequence 81, Appli
Sequence 41, Appli
Sequence 41, Appli
Sequence 41, Appli
Sequence 47, Appli
Sequence 49, Appli
Sequence 69, Appli
Sequence 69, Appli
Sequence 33, Appli
Sequence 67, Appli
Sequence 33, Appli
Sequence 33, Appli
Sequence 75, Appli
Sequence 75, Appli
Sequence 21, Appli
Sequence 63, Appli
Sequence 21, Appli
Sequence 63, Appli
Sequence 65, Appli
Sequence 77, Appli
Sequence 77, Appli
Sequence 43, Appli
Sequence 43, Appli
Sequence 7, Appli
Sequence 7, Appli
Sequence 45, Appli

NAME/KEY: CDS
LOCATION: (1)...(3633)
FEATURE:
NAME/KEY: misc_feature
LOCATION: (0)...(0)
OTHER INFORMATION: Cyl1218-2
US-11-058-727-3

Alignment Scores:
Pred. No.: 2,09e-64 Length: 3633
Score: 703.00 Matches: 219
Percent Similarity: 44.53% Conservative: 127
Best Local Similarity: 28.19% Mismatches: 289
Query Match: 18.17% Indels: 142
DB: 7 Gaps: 32

US-10-783-417-2 (1-735) x US-11-058-727-3 (1-3633)

QY 5 AsnAspAsnGluThrGluIleAlaSerHisThrSerProTyrPheProAsnArg 24
DB 10 AATAATCAAAATGAATATGAATATATAGATGCGACACCTTCT-----ACTTCT 57
QY 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
DB 58 GTATCCCAATGATCTTAACAGATACCTTTTTCGAATGAGCCAAATGCGCTACAAAAT 117
QY 44 ThrAsnTyrLysGluThrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
DB 118 ATGGATTATAAAGATTATTTAAAGATGCTGCGGGAATGCTAGTAGAATACCTCGTTCA 177
QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSer 82
DB 178 CCTGAAGTACTTGTAGCGGACAAGAT-----GCAGCTAAGCGCCCAATTGATATAGTA 231
QY 83 GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
DB 232 GGTAAATTAATCAAGCTTTAGG-----GTCCCATTTGTT 267
QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGlyGluGln 122
DB 268 GGGCCGATAGTGAGTCTTATCTACTCAACTATTGATATTCTGTGGCTTCAGGGCAA--- 324
QY 123 AspLysThrValThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu 142
DB 325 ---AAGAGTCAATGGAGATTTTATGGAACAAGTAGAAGAACTCATAAATCAAAAAATA 381
QY 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
DB 382 GCAGAAATATGCAAGGAATAAAGCGCTTTCCGGAATTAGAAGGATTAGGTAATAATTACCAA 441
QY 163 SerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
DB 442 TTATATCTAATCGCTTGAAGAAATGGAAGAA-----GTCCCATTTGTT 474
QY 183 ProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGluAsnVal 202
DB 475 AATCCAAATGGTTCA-----AGAGCGCTTACGAGATGTCGGAATCGAATTGAAATCCTG 528
QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
DB 529 GATAGTTTATTACGCAATACATGCCATCTTTTCGAGTGCACAAATTTTGAAGTACCATTTC 588
QY 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGlu 242
DB 589 CTTACAGTATATACACAGCGACCACTTCATTACTGTATTAAAGGACGCTTCAATT 648
QY 243 LeuAlaAspGluThrAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
DB 649 TTTGNGAAGATGGGATGG-----TCTCAACACCACT 681
QY 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
DB 682 ATTAATAACTATTATGATCGTCAATGAAATCTTACTGCAAGAAATATTCTGATCCTGTGTA 741

QY 283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTyrSerIle 302
DB 742 AAGTGGTATGAAACTGTTTAAAGAAATTAAGAGCGCACGCGCTAAACAATGGTGCAG 801
QY 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
DB 802 TATAACCAATTCGTAGAGAAATGACACTGACGGTTTTAGATGTTGTTCATATTATCCCA 861
QY 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIleLys 342
DB 862 AATTATGACACACGACGATAC-----CCAATGGAACGAAAGCA----- 900
QY 343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362
DB 901 ---CAACTAAACAAGGAAGTATATACA-----GATCCACTGGGCGCGTA 942
QY 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
DB 943 AAGCTGTCCTCAATTTGGTTCCTGTATGACAAAGCACCTTCTTCGGAGTGATAGAATCA 1002
QY 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
DB 1003 TCGGTTATTCGACCAACCC---CATGTATTGATTATATAACGGGACTCACAGTGTATACA 1059
QY 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408
DB 1060 CAATCAAGAAGCATTTCTCCGCTCGCTATATAAGACATTGGGCTGGTTCATCAATAAGC 1119
QY 409 AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySer 428
DB 1120 TATCATCGGATTTTGTAGTATAATATTATAAACAGATGTATGGAACCTAATAATCAATCTA 1179
QY 429 ProThrThrLysThrIleArgProPheGluSerTyr-----Lys 441
DB 1180 CACAGCACTAGTACCTTT---GATTTTACGAATATGATATTTACAGACGTTTATCAAAA 1236
QY 442 ValSerIleValThrArgArgGlnSerProProValSerProIleGlnProHisPheIle 461
DB 1237 GATGCGGTGCTCTTGATATTGTTTCTCGTTTATACGTATATA-----TTTTTTCGA 1290
QY 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
DB 1291 ATCCAGAGCTCGAGTCTTTCATGGTAAACCAATTTGAATTAATACAGAAAGACGTTAAAG 1350
QY 479 TyrSer-----AlaGlyGlySerLysSerAsnTyrGlnAsnThr 491
DB 1351 TATAATCCGGTTTCCAAAGATATTATAGCGGG----- 1383
QY 492 ThrPhePheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSer 511
DB 1384 -----ACAAGAGATTCGGAATTAGNAATTACCTCCAGAAACTTCA 1422
QY 512 -----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPhe---SerLeuPheThr 528
DB 1423 GATCAACCAAAATTTAGTTCATAGCATAGATTATGTATATCATATCACAAGTATTCCCGCG 1482
QY 529 TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThr 548
DB 1483 ACGGGTTTCAACTACCGGATTA-----GTACCTGTATTTTCTTGGACA 1524
QY 549 HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla 568
DB 1525 CATCGAGTCCCGATCTTATAAATGCGAGTTCATTCAGATAAAATTAATCAGATTCCGGTC 1584
QY 569 IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu 582
DB 1585 GTAAAGTTTCTGATTTGGCTCCCTCTATAACAGAGGCGGCAATAATAATACCGTTGTATCG 1644
QY 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGly----- 598
DB 1645 GGTCTCGATTATACAGGGGGGGGTAATAAAGTAATAAAGTAATGGAAGTAATTATATCA 1704


```
QY 599 -----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616
Db 1705 CATATCGGTGTTAAATTTCA-----GACATTAACAAGNATATAGTATGAGG 1752
QY 617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
Db 1753 ATTCGGTATGCTTCC-----GCTAATAATACCTGAAATTTTATATAAATCCTTCTCGAAGAA 1806
QY 637 GlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn 656
Db 1807 AAGCTTAATATCT---CACGCTCAAAAACCTATGAATAGAGGTGAAGCTTTTAAACATATAAT 1863
QY 657 AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValThr 673
Db 1864 AATATTAATATCGAGCTTGCCCTTAAATTTACGACACCGCACTTTCACTTACT 1923
QY 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693
Db 1924 CTAGGGGCT-----ATATTTGAAGCGGAGACACTTCTTGGAATTGAA 1965
QY 694 LeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGlu 713
Db 1966 GCTTATATAGCGAATGAAATTTATCCAGTAGATGAGACATAT-----GAAGCG 2016
QY 714 LysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsn 730
Db 2017 GAACAAGATTTAGACAGCGAAGAAAGCGTAGTGAATGCTTGTGTTACGAAT 2067
```

RESULT 2

```
US-11-108-389-3
; Sequence 3, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F. H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3
```

```
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3633)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Cry1218-2
```

US-11-108-389-3

```
Alignment Scores:
Pred. No.: 2,09e-64 Length: 3633
Score: 703.00 Matches: 219
Percent Similarity: 44.53% Conservative: 127
Best Local Similarity: 28.19% Mismatches: 289
```

```
Query Match: 18.17% Indels: 142
DB: 7 Gaps: 32
US-10-783-417-2 (1-735) x US-11-108-389-3 (1-3633)
QY 5 AsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyrPheProAsnArg 24
Db 10 AATAATCAAAATGAATATGAATATGAATATAGTCGACACCTTCT-----ACTTCT 57
QY 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
Db 58 GTATCCAATGATTTCAACAGATACCTTTTCCGAATGAGCCAAACAAATCGCTACAAAAT 117
QY 44 ThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
Db 118 ATGATTTATAAGATTAATTTAAATAATGTCGCGGAAATGCTAGTAATACCTCGTTCCTCA 177
QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIleValSer 82
Db 178 CCTGAAGTACTTGTAGCGGCAAGAT-----GCAGCTAAGCGCGCAATTGATATAGTA 231
QY 83 GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
Db 232 GGTAAATTAATCTATCAGCTTTAGGG-----GTCCCATTTGT 267
QY 103 GlyAlaIleIleIleSerPheGlyThrIleIleThrValPheTyrProAlaGlyGluGln 122
Db 268 GGGCCGATAGTAGTCTTTATCTCAACTTATTTGATATTTCTGTGGCTTCAGGGCAA--- 324
QY 123 AspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu 142
Db 325 ---AAGAGTCAATGGGAGATTTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATA 381
QY 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
Db 382 GCAGATATCAAGGAATAAAGCCCTTTCGGAATTAGAGGATTAGGTAATAATATACCAA 441
QY 163 SerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
Db 442 TTATATCTAACTGCGCTTGAAGATGGAAGAA-----474
QY 183 ProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGluAsnVal 202
Db 475 AATCCAAATGGTTC-----AGAGCCCTTACGAGATGTGCGAAATCGAATTTGAAATCCTG 528
QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
Db 529 GATAGTTTATTTAGCGCAATACATGCCATCTTTTCGAGTGACAAATTTTGAAGTACCATTC 588
QY 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGlu 242
Db 589 CTTACAGTATATACACAGCGAGCCCACTTCACTTCTGTTATTAAAGGACGCTTCAATT 648
QY 243 LeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
Db 649 TTTGGAGAAGAATGGGGATGG-----TCTACAACCACT 681
QY 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
Db 682 ATTAATAACTATTATGATCGTCAAAATGAACCTTACTGCAGAAATATTTGATCCTGTGTA 741
QY 283 AsnThrTyrArgThrGlyLeuLysAsnLeuAspGluProAsnMetLysTrpSerIle 302
Db 742 AAGTGTATGAAACTGCTTTAGCAAAATTAAGAGCGACGCGCTAAACAATGGGTGCGAC 801
QY 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db 802 TATAACCAATTCGCTAGAGAAATGACACTGACGCTTTTAGATGTTGTGCTATTATCCCA 861
QY 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyIleGluValLysGlyIleLys 342
Db 862 AATTATGACACACGACGCTAC-----CCAATGGAAACCAAGCA-----900
```

```
QY 343 AsnGluLeuThrArgGluLeuThrThrGluLeuAsnPheAspArgLeuProGlnLeu 362
Db 901 ----CACTAACAGGGAAGTATATACA-----GATCCACTGGCGCGGTA 942
QY 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
Db 943 AACGTGTCTCAATTCCTGGTATGACAAAGCACCTTCTTCGAGGTAGATCA 1002
QY 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheGlnGluGlnPheLeuThr 393
Db 1003 TCCGTTATTCGACCAACC-----CATGTATTGTATTATATAACAGGACTCAGGTATACA 1059
QY 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyLysSerAsnArg 408
Db 1060 CAATCAAGAAGCATTTCTTCGCTCGCTATATAAGACATTCGGCTGTCATCAATAAGC 1119
QY 409 AspAlaProThrTyrSerAsnThrThrGluThrLeuTyrGlyThrGlySer 428
Db 1120 TATCATCGGATTTTGTAGTGATTAATATATAAACAGATGATGAACTAATAAATCTA 1179
QY 429 ProThrThrLysThrLysArgProPheGluSerTyr-----Lys 441
Db 1180 CACAGCACTAGTACCTTT---GATTTACGAATTATGATATTTACAAGAGCTTATCAAA 1236
QY 442 ValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 461
Db 1237 GATCGGGTCTCTTGATATGTTTTCTCTGGTTATACGTATATA-----TTTTTGG 1290
QY 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
Db 1291 ATGCCAGAGTCGAGTGTTCATGTTGTAACCAATTGAATAATACCAGAAAGCGTTAAAG 1350
QY 479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr 491
Db 1351 TATAATCCGGTTCCAAAGATNTATAGCGGG-----1383
QY 492 ThrPhePheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCys 511
Db 1384 -----ACAAGAGATTCGAATTAGAAATTAACCTCCAGAAACTTCA 1422
QY 512 -----ProAsnPheAsnTyrSerHisIleLeuSerHisPhe---SerLeuPheThr 528
Db 1423 GATCAACCAAAATATGAGTCATATAGCATAGATTATGTATATCAATCAAGATATCCCG 1482
QY 529 TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyT 548
Db 1483 ACGGGTTCACACTACCGGATTA-----GTACCTGTATTTCTTCGGACA 1524
QY 549 HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIlePro 568
Db 1525 CATCGGAGTCCGATCTTATAAATGCGAGTTCATTCAGATAAATAATTAATCAGATTC 1584
QY 569 IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu 592
Db 1585 GTAAAGGTTTCTGATTTGGCTCCCTCTATAACAGGAGGCCAAATAATACCGTTGAT 1644
QY 593 GlyProGlyHisThrGlyAsnLeuValTyrLeuGlnSerGlnGly-----598
Db 1645 GGTCTGGATTTACAGGGGGGGGATAATAAAGTAATAAGAAATCGGATTAATATATCA 1704
QY 599 -----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIle 616
Db 1705 CATATCGGTGTTAAATTTCA-----GACATTAACAAGAATATAGTAGG 1752
QY 617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIle 636
Db 1753 ATTCCGTATGCTTCC-----GCTAATAATACTGAATTTTATATAAATCTCTCGA 1806
QY 637 GlyValIleGlyLeuProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn 656
Db 1807 AACGTTAAATCT---CACGCTCAAAAACTATGAATAGAGGTGAGCGTTTAAACATATA 1863
QY 657 AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValThr 673
```

```
Db 1864 AATTTAATTATCGCACTTTGCCCCCTATTAAATTTTACGACCAACCGAACCTTTCTACT 1923
QY 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693
Db 1924 CTAGGGGCT-----ATATTTCAAGCGGAGACTTCTTCTTGGATTGAA 1965
QY 694 LeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGlu 713
Db 1966 GCTTATATAGACCGAATCGAATTTATCCAGTAGATGAGACATAT-----GAAGCG 2016
QY 714 LysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsn 730
Db 2017 GAACAAGATTTAGACGACGCAAGCAAGACAGTGAATCGCTTGTTTACGAAT 2067

RESULT 3
US-11-058-727-18
; Sequence 18, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; PRIOR FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 6613
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Genomic Cry12I8-2
US-11-058-727-18

Alignment Scores:
Pred. No.: 5,54e-64 Length: 6613
Score: 703.00 Matches: 219
Percent Similarity: 44.53% Conservative: 127
Best Local Similarity: 28.19% Mismatches: 289
Query Match: 18.17% Indels: 142
DB: 7 Gaps: 32

US-10-783-417-2 (1-735) x US-11-058-727-18 (1-6613)
QY 5 AsnAspAsnAsnGluTyrGluIleAspSerHisThrSerProTyrPheProAsnArg 24
Db 1263 AATAATCAAAATCAATATGAATATGAATATATAGATGCGACACTTCT-----ACTTCT 1310
QY 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
Db 1311 GTATCAATGATTTCTAATACAGATACCTTTTGGCATGAGCCAAATATCGCTACAAAT 1370
QY 44 ThrAsnTyrLysGluTrpLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
Db 1371 ATGGATTATAAAGATTATTTAAAAATGCTCTCGCGGAAATGCTAGTGAATACCTGGTTCA 1430
```

QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIleValSer 82
Db 1431 CCGAGTACTTGTACGGACAAAGAT-----GCAGTAAAGCCGCAATGATAGTA 1484
QY 83 GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIle 102
Db 1485 GGTAAATTAATCACTACAGGTTAGG-----GTCCCAATTGTT 1520
QY 103 GlyAlaIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGln 122
Db 1521 GGGCCGATAGTACTTATCACTTATGATATTCGTCCTTGTGGCCTTCAGGGCAA--- 1577
QY 123 AspLysThrValTrpThrGlnPheIleLysMetClyGluIlePheValAspThrProLeu 142
Db 1578 ---AAGAGTCAATGGGAGATTTTATGGAAACAAGTAGAAGACTCAATAATCAAAAATA 1634
QY 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGlyPheArgGlnIleLeuGln 162
Db 1635 GCAGATATCGAAGGAATAAGCGCTTCGGAATTAGAGGATTAGGTAATAATTACAA 1694
QY 163 SerTyraThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
Db 1695 TTATATCTAATCGCTTGAAGATGAAAGAA----- 1727
QY 183 ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal 202
Db 1728 AATCCAAATGGTTCA-----AGAGCCTTACGAGATGTGCGAATCGATTGGAATCCTG 1781
QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrlsThrLeuLeu 222
Db 1782 GATAGTTATTAAGCAATACATCGCCATCTTCGAGTGACAAATTTTGAAGTACCATT 1841
QY 223 LeuProIleTyraGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGlyAlaGlu 242
Db 1842 CTTACAGTATATACACAGCGACCACTTCATTACTGTTTAAAGGACGCTCAATT 1901
QY 243 LeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
Db 1902 TTTGGAGAGAATGGGGATGG-----TCTCAACACCACT 1934
QY 263 SerAspAspTyrlsTyrlsLeuLeuLysGluAsnIleProLysTyrlsSerAsnTyrlsCysAla 282
Db 1935 ATTAATCAATTAATGATCGTCAAACTTACTGTCAGAAATATCTGATCATCTGTGA 1994
QY 283 AsnThrTyraArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle 302
Db 1995 AAGTGTATCAAACTGGTTTAGCAAAATTAAGGACGACGCGCTAAACAAATGGGTGCAC 2054
QY 303 PheAsnAspTyraArgTyrlsMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db 2055 TATAACCAATTCGCTAGAGAAATGACACTGACGGTTTATAGATGTTGTGCATTATCCCA 2114
QY 323 LeuTyraAspIleLysArgTyraAspSerIleGlyGlyIleGluValLysGlyIleLys 342
Db 2115 AATATATGACACACGACGTCAC-----CCAAATGGAAACGAAGCA----- 2153
QY 343 AsnGluLeuThrArgGluIleTyrlsThrGluIleAsnPheAspArgLeuProGlnLeu 362
Db 2154 ---CAACTAACAGGAGATATATACA-----GATCCACTGGGGCGGTA 2195
QY 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyrls 373
Db 2196 AAGCTGTCTCAATTTGTTCTCGTATGACAAAGCACCTTCTTCGGAGGATAGATACA 2255
QY 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrls 393
Db 2256 TCCGTTATTCGACACCC---CATGTATTTGATTATATATAACGGGACTCACAGTGTATACA 2312
QY 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408
Db 2313 CAATCAAGAGCAATTTCTCCGTCGTATATAGACATTTGGGCTGTCATCAAAATAAGC 2372

RESULT 4

QY 409 AspAlaProThrTyrlsSerAsnThrIleThrGluThrLeuTyrlsGlyGluArgThrGlySer 428
Db 2373 TATCATCGGATTTTATAGTATAATATAAAACAGATGTATGGAACATAATCAAAATCTA 2432
QY 429 ProThrTyrlsThrIleArgProPheGluSerTyrls-----Lys 441
Db 2433 CACAGCACTAGTACTT---GATTTTACGAATATATGATTTTACAAGACGTTATCAAAA 2489
QY 442 ValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIle 461
Db 2490 GATCGGTCCTCTGATATTTCTGTTTCTGTTTATACGTATATA-----TTTTTGA 2543
QY 462 IleAsnGlnIleGlyLeuTyrlsLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
Db 2544 ATGCCAGAAGTCGAGTTTTCATGTGTAACCAATTTGAATAATACCCAGAAAGACGTTAAAG 2603
QY 479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrlsGlnAsnThr 491
Db 2604 TATAATCCGGTTTCCAAAGATATATACGGGG----- 2636
QY 492 ThrPheGlnPheProArgLysAspCysAsnLeuValIleAspProGlyCysSer 511
Db 2637 -----ACAAGACATTCGGAATTAGAATTAACCTCCAGAAACTTCA 2675
QY 512 -----ProAsnPheAsnAsnTyrlsSerHisIleLeuSerHisPhe---SerLeuPheThr 528
Db 2676 GATCAACCAAAATTAATGATCATATAGCCATGATTTATGTCATATCAACAAGTATCCCGCG 2735
QY 529 TyrSerTyrlsValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTrpThr 548
Db 2736 ACGGTTCACTACCGGATTA-----GTACCTGATTTTCTTGGACA 2777
QY 549 HisSerValAspArgTyrlsAsnAlaIleSerAspLysIleIleThrMetIleProAla 568
Db 2778 CATCGAGTGCAGTCTTATAAATGACGTTCTATTCAGATAAAATTAATCAGATTCCGGTC 2837
QY 569 IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu 582
Db 2838 GTAAAGGTTTCTGATTTGGCTCCCTCTATAACAGAGGCGCAATAAATACCGTTGTATCG 2897
QY 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrlsLeuGlnSerGlnGly----- 598
Db 2898 GGTCTCGATTTACAGGGGGGGATATAAAGTAATAAGAAATGGAGTAATATATATCA 2957
QY 599 -----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrlsPheIleArg 616
Db 2958 CATATGCTGTAAATTTCA-----GACATTACAAGAATATATAGTATGAGG 3005
QY 617 LeuArgTyraAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
Db 3006 ATTCGGTATGCTTCC-----GCTAATAATACTGAATTTTATATAAATCTCTCGAAGAA 3059
QY 637 GlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrlsAsn 656
Db 3060 AAGCTTAAATCT---CACGCTCAAAAACCTATGAATAGAGGTGAAGCTTTTACATATAT 3116
QY 657 AsnLeuGlnTyrlsGlyAspPheGlyTyrlsPheGlnPheProSerThr-----ValThr 673
Db 3117 AAATTTAATTTATGCGACTTTGGCCCCCTTATTAATTTACGACCAACCGACCTTTCATTACT 3176
QY 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693
Db 3177 CTAGGGCT-----ATAATTGAAGCGGAAGACTTTTCTTGGAAATTGAA 3218
QY 694 LeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGlu 713
Db 3219 GCTTATATAGACCAATCGAATTTATCCAGTAGATGACATAT-----GAAAGC 3269
QY 714 LysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsn 730
Db 3270 GAACAGATTAGAACGACGAGAAAGACGTGAATGCTTGTGTTTACGAAT 3320

```

US-11-108-389-18
; Sequence 18, Application US/11108389
; Publication NO. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE REFERENCE: 357118/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 6613
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Genomic Cry1218-2
US-11-108-389-18

Alignment Scores:
Pred. No.: 5,54e-64 Length: 6613
Score: 703.00 Matches: 219
Percent Similarity: 44.53% Conservative: 127
Best Local Similarity: 28.19% Mismatches: 289
Query Match: 18.17% Indels: 142
DB: 7 Gaps: 32

US-10-783-417-2 (1-735) x US-11-108-389-18 (1-6613)
QY 5 AsnAspAsnGluTyrGluIleleAspSerHisThrSerProTyrPheProAsnArg 24
DB 1263 AATAATCAAAATGAATATGAAATTTATAGATGCGACACCTTCT-----ACTTCT 1310
QY 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
DB 1311 GPATCAATGATCTTAACAGATACCTTTTGGGAATGAGCCAAATGCGGTACAAAAT 1370
QY 44 ThrAsnTyrLysGluTyrPheLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
DB 1371 ATGGATTATAAGATTATTTAAATGTCTCGGGAATGCTAGTGAATACCTCGTTCA 1430
QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSer 82
DB 1431 CCTGAAGTACTTGTAGCGGACAAAGAT-----GCAGCTAAGGCGCAATTTGATATAGTA 1484
QY 83 GlyThrLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
DB 1485 GGTAATAATCTACTACGTTTAGGG-----GTCCCAATTTGTT 1520
QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGln 122
DB 1521 GGGCGATAGTAGCTTTTACTCAACTATTGATATTCTGTGGCTTCAGGGGAAA--- 1577
QY 123 AspLysThrValTrpThrGlnPheIleLysMetGlyIlePheValAspThrProLeu 142
DB 1578 ---AAGAGTCAATGGGAGATTTTATGGAAACAAGTAGAAGAACTCATATAATCAAAAATA 1634

```

```

QY 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
DB 1635 GCAGAATATCAAGGAATAAAGCGCTTTCGAATTTAGAGGATTAGGTAAATAATACCAA 1694
QY 163 SerTyrAsnThrAlaLeuAspAspTyrPheArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
DB 1695 TTATATCTAACTCGCTTGAAGAATGGAAGA----- 1727
QY 183 ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal 202
DB 1728 AATCCAAATGGTTCA-----AGAGCCTTAGGAGATGTGGAATTCGATTTGAATCTGTG 1781
QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
DB 1782 GATAGTTTATTACGCAATACATCCCATCTTTCGAGTGACAAATTTGAAGTACCATTTC 1841
QY 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnAlaGlu 242
DB 1842 CTTACAGTATATACAGGCGACCACTTCACTTACTGTATTATAAGGACGCTTCAATT 1901
QY 243 LeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
DB 1902 TTTGGAGAAGATGGGGATGG-----TCTACCAACCACT 1934
QY 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
DB 1935 ATTAATAACTATTATGATCGTCAATGAACCTTACTGCAGAAATATTCTGATCACTGTGA 1994
QY 283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle 302
DB 1995 AAGTGTGTATGAAACTGCTTTAGCAAAATTTAAAGGCGACGCGCTAAACAATGGTGCAC 2054
QY 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
DB 2055 TATAACCAATTCGTAGAGAAATGACACTGACGGTTTTAGATGTTGTGCAATTATCCCA 2114
QY 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyIleGluValLysGlyIleLys 342
DB 2115 AATTATGACACACGACGTAC-----CCAATGGAAACGMAAGCA----- 2153
QY 343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362
DB 2154 ---CAACTAAACAGGGAAGTATATACA-----GATCCACTGGGCGCGTA 2195
QY 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
DB 2196 AAGCTGTCTTCAATGGTTCCTGGTATGACAAAGCACCTTCTTTCGGAGTGATAGAATCA 2255
QY 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
DB 2256 TCGGTATTTCGACCAACC---CATGTATTGTATTATACGGGACTCACAGTGTATACA 2312
QY 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408
DB 2313 CAATCAAGAAGCATTTCTCCGCTCGCTATATAGACATTTGGGCTGCATCAATAAAGC 2372
QY 409 AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySer 428
DB 2373 TATCATCGGATTTTATGTAGTATAATATTATAAACAGATGATATGGAACCTAATAATCA 2432
QY 429 ProThrThrLysThrIleArgProPheGluSerTyr-----Lys 441
DB 2433 CACAGCACTAGTACCTTT---GATTTTACGAATTTATGATATTATACAGACGCTTATCAAA 2489
QY 442 ValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIle 461
DB 2490 GATGCGGTGCTCCTTGATATTGTTTCTCTGTTTATACGTATATA-----TTTTTTGGA 2543
QY 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrIleLys 478
DB 2544 ATGCCAGAGTCGAGTGTTCATGGTAAACCAATTTGAATAATACCAAGAAACGCTTAAG 2603

```

```
QY 479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr 491
|||...
Db 2604 TATAATCCGGTTTCCAAAGATATATAGCGGG-----2636

QY 492 ThrPhePheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSer 511
|||...
Db 2637 -----ACAAGAGATTCGGAATTAGAAATACCTCCAGAAACTTCA 2675

QY 512 -----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPhe---SerLeuPheThr 528
|||...
Db 2676 GATCAACCAAAATATGAGTCATATAGCATATGATCATATCACAAGATATTCGCGC 2735

QY 529 TyrSerTyrValIleGlyLeuGlnLeuIleAspThrGlyValLeuGlyTyrThr 548
|||...
Db 2736 ACGGTTTCAACTACCGAATTA-----GTACCTGATTTTCTTGACA 2777

QY 549 HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla 568
|||...
Db 2778 CATCGGAGTCCGATCTTATAATGCGATTCATTCAGATAAAATTAATCTCAGATTCGCGTC 2837

QY 569 IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu 582
|||...
Db 2838 GTAAGGTTTCTGATTTGGCTCCTCTATAACAGGAGGCGCAATAATACCGTTGTATCG 2897

QY 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGly-----598
|||...
Db 2898 GGTCTCGAATTTACAGGGGGGGGTAATAAAGTAATAGAAATGAGTAATATATATCA 2957

QY 599 -----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616
|||...
Db 2958 CATATCGGTTTAAATTTCA-----GACTAATACTACTGAATTTTATATAATCCTTCTGAAGAA 3005

QY 617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
|||...
Db 3006 ATTCCGTATGCTTCC-----GCTAATACTACTGAATTTTATATAATCCTTCTGAAGAA 3059

QY 637 GlyValIleGlyLeuProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn 656
|||...
Db 3060 AAGCTTAAATCT---CACGCTCAAAAAAATCATGAATGAGGTGAAGCTTTTAAACATATAAT 3116

QY 657 AsnLeuGluTyrClyAspPheGlyTyrPheGlnPheProSerThr-----ValThr 673
|||...
Db 3117 AAATTTAATATGCGACTTGGCCCTTATTAATTTACGACACCGAATCTTTCATTAAT 3176

QY 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaPheValSerAsnSerIle 693
|||...
Db 3177 CTAGGGGCT-----ATATTTGAAGCGGAAGACTTCTTGGAAATGAA 3218

QY 694 LeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGlu 713
|||...
Db 3219 GCTTATATAGACCGAATCGAATTTATCCAGTAGATGAGACATAT-----GAAGCG 3269

QY 714 LysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsn 730
|||...
Db 3270 GAACAAGATTTAGAACGAGCGAAGAAAGCAGTGAATGCTTGTTTACGAAT 3320

RESULT 5
US-11-091-643-5
; Sequence 5, Application US/11091643
; Publication No. US20050246789A1
; GENERAL INFORMATION:
; APPLICANT: TANAKA, Masao
; APPLICANT: YOKOYAMA, Tomoko
; APPLICANT: AOYAGI, Morichi
; APPLICANT: HASEGAWA, Makoto
; APPLICANT: EHARA, Gaku
; APPLICANT: KIMURA, Masaharu
; APPLICANT: NISHIHASHI, Hideji
; TITLE OF INVENTION: Polypeptide having larval growth inhibiting or
; TITLE OF INVENTION: insecticidal effect on scarabaeidae insects and
; TITLE OF INVENTION: polynucleotide encoding the same
; FILE REFERENCE: Op1335
; CURRENT APPLICATION NUMBER: US/11/091,643
```

```
; CURRENT FILING DATE: 2005-03-29
; PRIOR APPLICATION NUMBER: JP 2001-115754
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: JP 2001-203463
; PRIOR FILING DATE: 2001-07-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 4188
; TYPE: DNA
; ORGANISM: Bacillus popilliae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4158)
; US-11-091-643-5

Alignment Scores:
Pred. No.: 7.5e-63 Length: 4188
Score: 689.50 Match: 239
Percent Similarity: 45.07% Conservative: 113
Best Local Similarity: 30.60% Mismatches: 286
Query Match: 17.82% Indels: 143
DB: 9 Gaps: 38

US-10-783-417-2 (1-735) x US-11-091-643-5 (1-4188)

QY 1 MetAsnGln-----AsnAsnAspAsn-----AsnGluTyr 10
|||...
Db 55 ATGAATCAATATCATACCAAGAACGATAACAAAGTTTACAAACCAAGTGAAGTGAAGTG 114

QY 11 GluIleIle-----AspSerHisThrSerProTyrPheProAsnArgAsnSerAsnAsp 28
|||...
Db 115 CAAATATATACAACCTTCAAGTAACGCTTTACTTTACAGTCCCAAC-----159

QY 29 SerArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsnThrAsnTyrLysGlu 48
|||...
Db 160 ---AAGTATCCGATGCTACGATCCCAATGCTCATAGCAGAGGTAGAAATTTATAAAT 216

QY 49 TrpLeuAsnMetCysGlnGlyAsnThrGlnTyrGlyAspAsnPheGluThrPheAlaSer 68
|||...
Db 217 TGGTGGATATGTGTAGT-----GTAGGTGACGAT-----ACA 252

QY 69 AlaAspThrIleAlaAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAlaGly 88
|||...
Db 253 CGAAGTCCCGAAGTCCGGTGTACTGCACAAAGTTCTATTTCGACTTCTCTT-----303

QY 89 IleGlyClyLeuThrSer---IleSerGlyProIleGlyIle-----IleGlyAlaIle 105
|||...
Db 304 -----GGTATAACACGACACAAATCATTTGGCGCTCTAGGTATTCGCGTGTAGCGAAGCC 357

QY 106 IleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGlnAspLysThr 125
|||...
Db 358 ATTGGAATTTTCGCGCGCACTATTGGATTGGTGTGGCTCTGAGGTATTCGCGTGTAGTCCA-----411

QY 126 ValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeuThrGluSer 145
|||...
Db 412 ---TGGGTATATTTATGAATCATGTAGAGAACTCATATAATTCAAAATAACAGAGACT 468

QY 146 IleLysGlnLeuLysLeuGlnThrLeuGluClyPheArgGlnIleLeuGlnSerTyrAsn 165
|||...
Db 469 GTAAAAAATGAGGCAATTAACAAGATTAGACGGTTTATAGCAATGTCTAGCGCTATATCAA 528

QY 166 ThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeuProSer 185
|||...
Db 529 AAGGCTTTCGAAGAATGCGAA-----CAACATCCT-----558

QY 186 SerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAsp 205
|||...
Db 559 ---ACTCTTGAATCAGCTCCGACTACGCTGTAACAGATGATTTTCTAATGTAATAATTT 615

QY 206 PheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeuLeuProIle 225
|||...
Db 616 TTTGAGGCATTTATGCTTTCATTTCAGAGTACCAGGTTATGAGTACCATTATTAAGCGTA 675
```

QY 226 TyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAsp 245
DB TAGTATCCGCTCGCAACCTCCATTTATTATTATTAAAGAGATAGCTCGATTTTCGGTTG 735
QY 246 GluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAsp 265
DB GAITGGGGATTA-----AGTCAACATCATGTTAAAC-----GAT 768
QY 266 TyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyr 285
DB AATTATAATCTCCAATAGGCGCTCGCAGATTATGCAAAATCATTTGTACAACTTGGTAT 828
QY 286 ArgThrGlyLeuLysAsnLeuArgAspGluProAsnMethIysTrpSerIlePheAsnAsp 305
DB CGGACGGGTTTACAAAGATTGCAAGGCCAACATGTAGCAGTTGGGTCAATTATATATCGA 888
QY 306 TyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyrAsp 325
DB TTTAGAGAGAAATGACACTAATCTGTATTAGATGTTTGGCATATTATTTCAAGTTATGAT 948
QY 326 IleLysArgTyrArgAspSerIleGlyIleGluValLysGlyIleLysAsnGluLeu 345
DB TATCGTAGTTAC-----CCAAATGGAGCTAAGGGA-----GAGCTT 984
QY 346 ThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeuArgValGln 365
DB ACCGAGAGAAATTTATACGATCCAGTACGAGGCTCTTTT-----TGGGTGAATAGAGCA 1038
QY 366 ProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPheLysLeuPheSerPhe 385
DB CCAAACTTCGCATCAATGAAAATACAGTAGTTAGGCAACACACACCC-----TTACTTGG 1095
QY 386 LeuGluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIle 405
DB CTAGTTACTTTTAACTTAATACAGTCAAGTCAGATCTGGCGAT-----GGAAAT 1146
QY 406 SerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArg 425
DB TCTAATATTATGGAAATACACATAGTCAACCGTCAAGTGAACCGGAGGTCAGTCCCT 1206
QY 426 ThrGlySerProThr-----ThrLysThrIleArgProPheGluSerTyrLysVal 442
DB ATTCAGAGCTCTACCTGTGGAAGTACTGTGTACAAAT-----TATCGC--- 1248
QY 443 SerIleValThrAspArgGln-----SerProValSerProIleGlnProHisPhe 460
DB -----ACGATAAATTTACTTTTAAATCCA-----TTT 1275
QY 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr 476
DB TTATTAGGTGATATTATTATACCATTAATACAGGTATTGTTCTTATCTGCGTAATTTGTTT 1335
QY 477 LeuLysTyrSerAlaGlyGlySerLeuSerAsn-----TyrGlnAsnThr 491
DB GGAATCTATTACAGTAGATTACGACGACTCGTGTATTGAGCTTCTGTATGAGAACCA 1395
QY 492 ThrPheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSer 511
DB AGAGTTTTCGCCGCTTACAAATCATCAAAATTCGTGAATTA-----CCTGAGGTAAAC 1446
QY 512 -----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThr 528
DB TCGGATAGCCCAACTGCTCCGCACTATGATCATAGACTATCGTATATCTCAGGTTTTCGA 1506
QY 529 TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGly-----Val 544
DB -----ACTGATGTGGGAGAACGGTTCTAGTT 1533
QY 545 LeuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSer---AspLysIleIle 563
DB TATGGGTGGACATCTTCAACTGCTACTCGTGAAGAAATAATATTACGTAGACGACAGATA 1593

QY 564 ThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGluGly 583
DB GTACAACTTCCAGCTGTTAAGGGAACAAGTCTCAAC---AATTGCCAAGTAGTTAGAGGA 1650
QY 584 ProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIleThr 603
DB ACTGGATTTTACAGGAGGAGACTGGTTGAAGCCTTAATATATGTTTCTCTAGCT 1710
QY 604 CysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGly 623
DB CTTCGGTTTC---AGTTCGACTTACACTTACCGCTCCGCATTCTGTATGCTGCCGACGA 1767
QY 624 AlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIleProPro 643
DB GCGCGATCA-----GGTTTTCTCTGTATATATCGGATCAATATGAGAAATTTTCCA 1818
QY 644 GlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn-----AsnLeuGlnTyr 660
DB ---ACCACACAGATATCGCTTCTCCACAAATGTACTCTGCCCCAAAATGTACCATAC 1875
QY 661 GlyAspPheGlyTyrPheGlnPheProSerThrValThrLeu----- 674
DB GAGGCTTTTAAGATTGTAGATTTACCTTCTACTGTACTATTAGAAATCTTCTCTCTGCT 1935
QY 675 -----ProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSer 692
DB TCAACAACTTTTCGACTTGTATTTCCGTTTTCATTTGTCATTAGGAATATCTGCAAAAT--- 1992
QY 693 IleLeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArg 712
DB ---ATATTAAATGACCGAATTTGAATTTGTTCCCATGAGGGTTCCTTGTTCGAGTAGCAA 2049
QY 713 GluLysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsnHisThr 732
DB ACCAAACAGCAGCTAGAAAAGCAAGGCGGTGAACCATTTGTTTACAGATGGATCG 2109
QY 733 Lys 733
DB 2110 AAA 2112
RESULT 6
US-11-058-727-13
; Sequence 13, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 2022
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (truncated)
; FEATURE:
; NAME/KEY: CDS


```
/ LOCATION: (1)...(2022)
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (0)...(0)
/ OTHER INFORMATION: 1218-2A
/ US-11-058-727-13

Alignment Scores:
Pred. No.: 3 52e-62      Length: 2022
Score: 678.50           Matches: 212
Percent Similarity: 44.61%      Conservative: 123
Best Local Similarity: 28.23%    Mismatches: 277
Query Match: 17.54%           Indels: 139
DB: 7                     Gaps: 31

US-10-783-417-2 (1-735) x US-11-058-727-13 (1-2022)

QY 5 AsnAspAsnGluThrGlyLeuValSerHisThrSerProTyrPheProAsnArg 24
DB 10 AATAATCAAAATGAATATGAATATATAGATGCGACACTTCT-----ACTTCT 57

QY 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
DB 58 GTATCCAAATGATTTAACAAGATCCCTTTTGGCAATGAGCCCAACAAATGCGCTACAAAAT 117

QY 44 ThrAsnTyrLysGluThrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
DB 118 ATGGATTATAAAGATATTATTAATAATGTCGCGGAAATGCTAGTGAATACCCCTGGTTCA 177

QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSer 82
DB 178 CCTGAAGTACTTGTAGCGGACAAGAT-----GCAGCTAAGCGCGCAANTGATATAGTA 231

QY 83 GlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGlyIle 102
DB 232 CGTAATATTACTACAGTTTAGG-----GTTCCCATTTGTT 267

QY 103 GlyAlaIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGln 122
DB 268 GGGCCGATAGTAGTCTTTATATCTCAACTTATTGATATTCTGTGGCTTCAGGCGCA--- 324

QY 123 AspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu 142
DB 325 ---AAGAGTCATGGAGAGATTTTATGGAAACAAGTAGAAGAACTCAATAATCAAAAATA 381

QY 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
DB 382 GCAGATATGCAAGGAATAAAGCGCTTTCGGAATTAGAGGATTAGTATATAATTACCAA 441

QY 163 SerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
DB 442 TTATATCTAACTCGCTTGAAGAATGMAAGAA----- 474

QY 183 ProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGluAsnVal 202
DB 475 AATCCAAATGGTTCA-----AGAGCCTTACGAGATGTGCGAAATGATTTGAATCCTG 528

QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
DB 529 GATAGTTTATTACGCAATACATGCCACTCTTTTCGAGTGCACAAATTTTGAAGTACCATT 588

QY 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGlyAlaGlu 242
DB 589 CTTACAGTATATACACAGCAGCAACCTTCATTACTGTATTAAAGGACGCTTCAATT 648

QY 243 LeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
DB 649 TTTGAGAGAAGATGGGATCG-----TCTCAACACCACT 681

QY 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
DB 682 ATTAATAACTATTATGATCGTCAAAATGAAACTTACTTGCAGAAATATTCTGATCCTGTGA 741
```

```
QY 283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle 302
DB 742 AAGTGGTATGAACACTGGTTTAGCAAAATTAAGAGCCAGAGCGCTAAACAATGGGTCGAC 801

QY 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
DB 802 TATAACCAATTCCTAGAGAAATGACACTGACGGTTTGTAGATGTTGCTGCAATTATCCCA 861

QY 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyLeuValLysGlyIleLys 342
DB 862 AATTATGACACACGACGCTAC-----CCAATGGAACCAAGCA----- 900

QY 343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362
DB 901 ---CAACTAACAGGAAGTATATACA-----GATCCACTGGGCGGTA 942

QY 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
DB 943 AACGTGCTTCAATGGTTCCTGGTATGACAAAGCACCTTCTTTGGAGTGATAGATCA 1002

QY 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
DB 1003 TCCGTATTTCGACCACCC---CATGTATTGATTATATAACGGGACTCACAGTGATACA 1059

QY 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408
DB 1060 CAATCAAGAGACATTTCTCCGCTGCTATATAAGACATTTGGGCTGCATCAATAAAG 1119

QY 409 AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySer 428
DB 1120 TATCATCGGATTTTGTAGTAAATATTATAAACAGATGTATGGAACCTAATAAATCTA 1179

QY 429 ProThrThrLysThrIleArgProPheGluSerTyr-----Lys 441
DB 1180 CACAGCACTAGTACTTT---GATTTTACGAATTATGATATTATCAAGACGTTATCAAAA 1236

QY 442 ValSerIleValThrAspArgGlnSerProValSerProIleGlnProHisPheIle 461
DB 1237 GATCGGTGCTCTGTGATATTGTTTTCCTGGTTATACGTATATA-----TTTTTGA 1290

QY 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
DB 1291 ATCCAGAGTCGAGTGTTCATCGTAAACCAATTTGAATAATACCAGAAAGACGTTAAAG 1350

QY 479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr 491
DB 1351 TATAATCCGGTTTCCAAAGATATTATAGCGGG----- 1383

QY 492 ThrPhePheGlnPheProArgLysAspCysAsnLeuValIleAspProGlyCysSer 511
DB 1384 -----ACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCA 1422

QY 512 -----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPhe---SerLeuPheThr 528
DB 1423 GATCAACCAAAATATTAGTCATATAGCCATAGATTATGTCATATACAGATATTCCCGCG 1482

QY 529 TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTrpThr 548
DB 1483 ACGGTTCAACTACCGGATTA-----GTACCTGTATTTTCTTGGACA 1524

QY 549 HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla 568
DB 1525 CATCGGAGTGCCGATCTTATAAATGCGAGTTTCATTAGATAAAATTTACTCAGATTCCGG 1584

QY 569 IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu 582
DB 1585 GTAAAGGTTCTGATTTGGCTCCCTCTATAACAGAGGGCCCAATAATACCTGTGTATCG 1644

QY 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGly----- 598
DB 1645 GGTCCTGGATTTACAGGGGGGGGATAATAAAGTAATAAGAAATGGAGTATTAATATCA 1704

QY 599 -----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616
```


Db	943		AACGTGCTTCAATGGTTCCTGGTATGACAAAGACACCTTCTTCGGAGTGATAGATCA	1002
Qy	374	AsnLeuThrArgAlaSerPheLeuLysPheLeuSerPheLeuGluGlnPheIlePheTyrThr	393	
Db	1003	TCGGTTATTCGACACCC--CATGTATTGATATATATAACGGGACCTCACAGTGTATACA	1059	
Qy	394	GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg	408	
Db	1060	CAATCAAGAAAGCATTTCTCCGCTGCTATATAAGACATTTGGGTGCTCATCAATAAGC	1119	
Qy	409	AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySer	428	
Db	1120	TATCATCGGATTTTGTAGTAAATATATAAACAGATGTATGAACTAATCAAAATCTA	1179	
Qy	429	ProThrThrLysThrIleArgProPheGluSerTyr-----Lys	441	
Db	1180	CACAGCACTAGTACCTTT--GATTTTACGAATTATGATATTTACAAGAGCTTATCAAA	1236	
Qy	442	ValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIle	461	
Db	1237	GATCGGTGCTCTCGATATGTGTTTTTCTCGTTATACGTATATA-----TTTTTGG	1290	
Qy	462	IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys	478	
Db	1291	ATGCCAGAAGTCAGTTTTTTCATGGTAAACCAATTTGAATTAATACCAGAAAGACGTTAAAG	1350	
Qy	479	TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr	491	
Db	1351	TATAATCCGGTTTCCAAAGATATATATAGCGGG-----	1383	
Qy	492	ThrPheGlnPheProArgLysAspCysAsnLeuValIleAspProGlyCysSer	511	
Db	1384	-----ACAAGAGATTCGGAATTAGAAATTACCTCCAGAAACTTCA	1422	
Qy	512	-----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPhe--SerLeuPheThr	528	
Db	1423	GATCAACCAAAATATGAGTCATATAGCCATAGATTTATGTCATATACAAGATATTC	1482	
Qy	529	TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThr	548	
Db	1483	ACGGGTCACTACCGGATTA-----GTACTGTATTTCTTGACACA	1524	
Qy	549	HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla	568	
Db	1525	CATCGGAGTCCGATCTTATAAATGCAGTTTCATTCAGATAAATTTACTCAGATTC	1584	
Qy	569	IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu	582	
Db	1585	GTAAGAGTTTCTGATTTGGTCCCTCTATTAACAGGAGGGGCAAAATAATACCGTTGTATCG	1644	
Qy	583	GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGly-----	598	
Db	1645	GGTCCTGGATTTACAGGGGGGGGATATAAAGTATATAAGAAATGAGTAATATATCA	1704	
Qy	599	-----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg	616	
Db	1705	CATATGCGTGTAAATTTTCA-----GACATTAACAAAGAATATATGATGAGG	1752	
Qy	617	LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro	636	
Db	1753	ATTCCGTATGCTTCC-----GCTAATAATTAATGAAATTTTATATAAATCCTCTCGAGAA	1806	
Qy	637	GlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn	656	
Db	1807	AACGTTAAATCT--CACGCTCAAAAACATGATGAATAGAGGTGAAGCTTTAAACATATAAT	1863	
Qy	657	AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValThr	673	
Db	1864	AAATTTAATTATCGGACTTTGCCCTTATTAATTTACGACACCGNACCTTTCATTACT	1923	
Qy	674	LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaIleAspValSerAsnSerIle	693	

Db 232 GGTAAATTACTACAGTTTAGGG-----GTCCCAATTGTT 267
Qy 103 GlyAlaIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGln 122
Db 268 GGGCCGATAGTAGTCTTTATATCACTTATGATATTCTGTGGCTTCAGGGAA--- 324
Qy 123 AspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu 142
Db 325 ---AAGAGTCAATGGGAAATTTTATGGAACAAGTAGAAGAACTCAATTAATCAAAAATA 381
Qy 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
Db 382 GCAGATATGCAAGSAATAAAGCGCTTCGGAATTAGAGGATTAGGTAATAATACCAA 441
Qy 163 SerTyrAsnThrAlaLeuAspAspTTPArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
Db 442 TTATATCTAACTCGCTTGAGAAATGGGAAGAA----- 474
Qy 183 ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal 202
Db 475 AATCCAAATGGTTCA-----AGAGCCTTACGAGATGTGGAAATCGATTTGAAATCCTG 528
Qy 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
Db 529 GATAGTTATTTACGCAATATATGCCATCTTTTAGAGTGAACAATTTTGAAGTACCATT 588
Qy 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlyAlaGlu 242
Db 589 CTTACTGTATGCAATGGCAGCACTTCACTTACTGTATTAAAGACGGTCAATT 648
Qy 243 LeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
Db 649 TTTGGAGAAGATGGGGATCG-----TCAACAACACTACT 681
Qy 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
Db 682 ATTAATAACTATTATGATCGTCAATGAACTTACTGCAGAAATATCTGATCACTGTGTA 741
Qy 283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle 302
Db 742 AAGTGTATGAAACTGGTTTTAGCAAAATTAAGGCGACGCGCTAAACAATGGGTGAC 801
Qy 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db 802 TATAACCAATTCGTAGAGAAATGACACTGGCGTTTGTAGATGTGTGCAATTATCCCA 861
Qy 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyLeuValLysGlyIleLys 342
Db 862 AATTATGACACGACGCTAC-----CCAAATGGAAACGAAAGCA----- 900
Qy 343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362
Db 901 ---CAAATAACAGGGAAGTATATACA-----GATCCACTGGCGCGGTA 942
Qy 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
Db 943 AAGCTGCTTCAATTGGTCTCGTATGACAAAGCACTTCTTCGGAGTGAATATCA 1002
Qy 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
Db 1003 TCGGTTATTCGACCAACC---CATGTATTTGATTATATACGGGACTCAGAGTGTATACA 1059
Qy 394 GluAsnThrAsnPheGlyAsn---ArgLeuVal-----GlyIleSer 406
Db 1060 CAATCAAGAGCAATTCTTCGCTCGCTATATAGACATTTGGGCTGCTCAATAAAGC 1119
Qy 407 AsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThr 426
Db 1120 TACCATCGTGTAGTAGGGGTAGTAAT-----CTTCAACAATGATAGAACTAATCAA 1173
Qy 427 GlySerProThrThrLysThrIleArgProPheGluSerTyrLysValSerIleValThr 446

Db 1174 AATCTACACAGCACTAGTACCTTT---GATTTTACGAATTATGATATTTACAAGACTCTA 1230
Qy 447 AspArgGlnSerProProValSerProIleGlnPro-----HisPheIle 461
Db 1231 TCAAGAGTACAGTACTCTTGATATTGTTTACCTCGTATATACGTATATATTTTTCGA 1290
Qy 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
Db 1291 ATGCCAAGTCCGAGTTCATCGTTAAACCAATTTGAATAATACAGAAAGACGTTAAAG 1350
Qy 479 TyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArg 498
Db 1351 TAT-----AATCCAGTTTCCAAAGATATATATACGAGTACAAGA 1389
Qy 499 LysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSer 518
Db 1390 GATTCGGAATTAGNATTACCTCCAGAAACTTCAGATCAACCAATTTATGAGTCATATAGC 1449
Qy 519 HisIleLeuSerHisPhe---SerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeu 537
Db 1450 CATAGATTATGTCTATATCACAGTATTCCCGCGACGGTAACACTACCGGATTA----- 1503
Qy 538 GlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAla 557
Db 1504 -----GTACCTGTATTTCTTGGACACATCGAAGTCGAGATTTAAACAATAACA 1551
Qy 558 IleSerAspLysIleIleThrMetIleProAlaIleLys---GlyAsnAsnLeuAspThr 576
Db 1552 ATATATTCAGATAAAATCACTCAAAATTCGCGCGTTAAATGTTGGGATAAATTACCGTTT 1611
Qy 577 AsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuVal---TyrLeuGln 595
Db 1612 ---GTTCCAGTGGTAAAGACGAGACATACAGAGGGGATTTATTACAGTATAATAGA 1668
Qy 596 SerGlnGly-----ArgLeuGluIleThrCysGluThrPro 607
Db 1669 AGTACTGCTCTGTAGGAACCTTATTCTAGTCGATATGCGCTAGCATTAGAA----- 1722
Qy 608 AsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThr 627
Db 1723 ---AAGACGGGAATATCGTGAAGACTGAGATATGCTACTGATGCA----- 1767
Qy 628 LeuProAsnIleSerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsn 647
Db 1768 -----GATATTGTATTGATGTAACGATCTCAGATTCAGATGTCACCAAAACAATGAAC 1821
Qy 648 AsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGln 667
Db 1822 CCAGGT-----GAGGATCTGCATCTAAACCTTTTAAAGTTGCAGAT 1863
Qy 668 PheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAla 687
Db 1864 GCTATCACACATTAATAATTAGCAACAGATGTCGTAGCAITTTGAACAATAAATTAGGT 1923
Qy 688 AspValSerAsnSer-----IleIleIleIleAspLysIleGluPheIlePro 703
Db 1924 GAAGACCTTAATCAACATTTATCTGTATAGTTTACGTTGACCGAATCAATTCATCCCA 1983
Qy 704 IleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLys 723
Db 1984 GTAGATGAGACATAT-----GAAGCGGAACAAGATTTTAGAAGCAGCGAAGAAAGCA 2034
Qy 724 IleAsnThrPhePheThrAsn 730
Db 2035 GTGAATCCCTTGTTTACGAAT 2055

RESULT 9

US-11-108-389-1
; Sequence 1, Application US/11108389
; Publication NO. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan

QY	5	AsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyrPheProAsnArg	24
Db	10	ATAAATCAAAATGAATATGAAATTATAGATCGCACCCCTCT	57
QY	25	AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGluAsn	43
Db	58	GTATCCAATGATTTCTAAACAGATACCCCTTTTCGGAATGAGCCACAAATGGCGCTACAAAAT	117
QY	44	ThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn	62
Db	118	ATGGATATTAAGAGATATTTAAAAATGCTCGGGGAATGCTAGTAATACCCCTGGTTC	177
QY	63	PheGluThrPheAlaSerAlaAspThrIleAlaIaValSerAlaGlyThrIleValSer	82
Db	178	CCTGAAGTACTTGTAGCGGCAAGAT-----GCAGCTAAGGCGCAATTGATATAGTA	231
QY	83	GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIle	102
Db	232	CGTAAATTACTATCAGGTTTAGG-----GTCCCAATTGTGT	267
QY	103	GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGlyGluGln	122
Db	268	GGGCGCATAGTCAGGCTTTATCTACTCAACTTATGATATTCGTGGCCCTTCAGGGAA--	324
QY	123	AspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu	142
Db	325	---AGAGTCAATGGGAATTTTATGGAAACAGTAGAAGACTCATATATCAAAAAATA	381
QY	143	ThrGluSerIleLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln	162

QY 499 LysLysAspCysAsnLeuValIleAAspProGlyCysSerProAsnPheAsnAsnTyrSer 518
Db 1390 GATTTCGAATTAGAATACCTCCAGAACTTCAGATCAACCAAAATATAGTCATATAGC 1449
QY 519 HisIleLeuSerHisPhe---SerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeu 537
Db 1450 CATAGATTATGTCATATCAAGATTTCGCGACGGGTAACTACCTACCGGATTA----- 1503
QY 538 GlnIleLeuAspThrGlyValLeuGlyTyrHisSerSerValAspArgTyrAsnAla 557
Db 1504 -----GTACCTGTATTTCCTTGACACATCGAAGTCAGATTAAACATACA 1551
QY 558 IleSerAspLysIleIleThrMetIleProAlaIleLys---GlyAsnAsnLeuAspThr 576
Db 1552 ATATATTTCAGATAAAATCACTCAAAATCCGCGCGTTAAATGTTGGGATAATTACCGTTT 1611
QY 577 AsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuVal---TyrLeuGln 595
Db 1612 ---GTTCCAGTGGTAAAGGACGAGGACATACAGGAGGGGATTTATTACAGTATAATAGA 1668
QY 596 SerGlnGly-----ArgLeuGluIleThrCysGluThrPro 607
Db 1669 AGTACTGGTCTCTAGGAACCTTATTCTAGCTCGATATGCGCTAGCAATTAGAA----- 1722
QY 608 AsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThr 627
Db 1723 ---AAAGCAGGGAATATCGTGAAGACTGAGATATGCTACTATGCA----- 1767
QY 628 LeuProAsnIleSerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsn 647
Db 1768 -----GATATTGTAATTGATGTAAACGATGCTCAGATTGAGTGCACAAACAAATGAAC 1821
QY 648 AsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGln 667
Db 1822 CCAGGT-----GAGGATCTCAGCATCTAAAACTTTTAAAGTTGCAGAT 1863
QY 668 PheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAla 687
Db 1864 GCTATCACACATTAATTTAGCAACAGATAGTTCGTAGCATTTGAACATATTTAGGT 1923
QY 688 AspValSerAsnSer-----IleLeuIleIleAspLysIleGluPheIlePro 703
Db 1924 GAAGACCTAAATCAACATTTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCA 1983
QY 704 IleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLys 723
Db 1984 GTAGATGAGACATAT-----GAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCA 2034
QY 724 IleAsnThrPhePheThrAsn 730
Db 2035 GTGAATGCCTTGTGTACGAAT 2055

RESULT 10

US-11-058-727-17
; Sequence 17, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE REFERENCE: 35718/287809
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786

; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 4874
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Genomic DNA 1218-1
US-11-058-727-17

Alignment Scores:

Pred. No.:	3,51e-61	Length:	4874
Score:	675.00	Matches:	216
Percent Similarity:	44.46%	Conservative:	125
Best Local Similarity:	28.16%	Mismatches:	300
Query Match:	17.45%	Indels:	126
DB:	7	Gaps:	32

US-10-783-417-2 (1-735) x US-11-058-727-17 (1-4874)

QY 5 AsnAspAsnAsnGluTyrGluIleAspSerHisThrSerProTyrPheProAsnArg 24
Db 740 AATATCAAAATGAATATGAATATATAGATCGGACACCTTCT-----ACTTCT 787
QY 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
Db 788 GTATCCATGATCTTAACACATACCTTTTGGGATGAGCCACAAATCGCTACAAAT 847
QY 44 ThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
Db 848 ATGGATTATAAGATTATTTAAAAATGCTCGCGGAAATGCTAGTGAATACCTGGTTCA 907
QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSer 82
Db 908 CTGAAGTACTGTGTAGCGGACAAAGAT-----GCAGCTAAGCGCCCAATGATATAGTA 961
QY 83 GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
Db 962 GGTAAATTACTATCAGGTTTAGGG-----GTCCCATTTGTT 997
QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGlyGluGln 122
Db 998 GGGCCGATAGTAGTCTTTATCTCAACTTATTGATATCTGTGGCTTCAGGGGAA--- 1054
QY 123 AspLysThrValThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu 142
Db 1055 ---AAGACTCAATGGGAAATTTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATA 1111
QY 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGlyGlyPheArgGlnIleLeuGln 162
Db 1112 GCAGATATCAAGGAATAAGCGCTTCGGAATTAGAGGATTAGGTAATAATTACCAG 1171
QY 163 SerTyrAsnThrAlaLeuAspAspTyrPArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
Db 1172 TTATATCTAATCGCTTGAAGAATGGGAAGAA----- 1204
QY 183 ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal 202
Db 1205 AATCCAAATGGTTCA-----AGAGCCTTCAGAGATGCGAAATCGAATTGAAATCCTG 1258
QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
Db 1259 GATAGTTTATTACGCAATATATGCCATCTTTTAGAGTGAACAATTTGAAGTACCATTTC 1318
QY 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGlu 242

1319	Db	CTTACTGTATATGCAATGGCAGCCACCTTCATTACTTACTGTATTATAAGGACCGCTCAATT	1379
243	Qy	LeuAlaAspGluThrProAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr	262
1379	Db	TTTGGAGAAGAATGGGGATGG-----TCAACAACACTACT	1411
263	Qy	SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla	282
1412	Db	ATTAAATAACTATTATGATCGTCAAATGAACACTTACTCGAGAATATTCTGATCACTGTGTA	1471
283	Qy	AsnThrTyrArgThrGlyLysLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle	302
1472	Db	AAGTGGTATGAACTGGTTTAGCAAAATTAAGCGCAGGCGCTAAACAATGGGGTTGAC	1531
303	Qy	PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer	322
1532	Db	TATAACCAATTCCTAGAGAAATGACACTGGCGGTTTTAGATGTGTTCATTATTTCCCA	1591
323	Qy	LeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIleLys	342
1592	Db	AATTATGACACACGACCGCTAC-----CCAATGGAAACGAAGAACA-----	1630
343	Qy	AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu	362
1631	Db	---CAACTAACACGGAAGTATATACA-----GATCCACTGGGGCGCGTGA	1672
363	Qy	ArgValGln-----ProAsnLeuAlaThrMetGluTyr	373
1673	Db	AACGTGCTTCAATTGGTTCTCGTATGACAAAGCACCTTCTTCGGAGTGTAGATCA	1732
374	Qy	AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr	393
1733	Db	TCGGTATTTCGACCACCC---CATGTATTTCATTATATAACGGGACTCACAGTGTATACA	1789
394	Qy	GluAsnThrAsnPheGlyAsn---ArgLeuVal-----GlyIleSer	406
1790	Db	CAATCAAGAAGCATTTCTCCGCTCGTATATAGACATTGGGCTGTCATCAAAATAGC	1849
407	Qy	AsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyValArgThr	426
1850	Db	TACCATCGTGTAGTAGGGTAGTAAT-----CTTCAACAATGTATAGACTAATCA	1903
427	Qy	GlySerProThrThrLysThrIleArgProPheGluSerTyrLysValSerIleValThr	446
1904	Db	AATCTACACAGCAGCTAGTACTTT---GATTTTACGAATTATGATATTTCACAGACTCTA	1960
447	Qy	AspArgGlnSerProProValSerProIleGlnPro-----HisPheIle	461
1961	Db	TCAAAGATGACGACTCTCTTGATTTGTATCCCTGTTATACGTATATATTTTGGGA	2020
462	Qy	IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys	478
2021	Db	ATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTGAATATATACCAGAAAGAGCTTAAG	2080
479	Qy	TyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArg	498
2081	Db	TAT-----AATCCAGTTTCCAAAGATATTATAGCGAGTACAAGA	2119
499	Qy	LysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSer	518
2120	Db	GATTCGGAAATTAGAAATTTACCTCCAGAAACTTCAGATCAACCAATTTATGATCATATAGC	2179
519	Qy	HisIleLeuSerHisPhe---SerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeu	537
2180	Db	CATAGATTATGTCATATACAAAGTATTTCCTCCGCGACGGGTAACTACCGGATTA-----	2233
538	Qy	GlnIleLeuAspThrGlyValLeuGlyThrHisSerSerValAspArgTyrAsnAla	557
2234	Db	-----GTACCTGTATTTCTTGGACACATCGAAGTGCAGATTATAACNATACA	2281
558	Qy	IleSerAspLysIleIleThrMetIleProAlaIleLys---GlyAsnAsnLeuAspThr	576
2282	Db	ATATATTTCAGATAAAATCACTCAAATTCGCGCGTTAAATGTTGGGATATATTTACCGTTT	2341

```

Qy 577 AsnSerIysValIleGluGlyProGlyHisThrGlyGlyAsnLeuVal-----TyrLeuGln 595
Db 2342 ---GTTCCAGTGGTAAAGAACCCAGGACATACACAGGAGGGATTTATTTCACAGTATAATAGA 2398
Qy 596 SerGlnGly-----ArgLeuGluIleThrCysGluThrPro 607
Db 2399 AGTACTGGTTCTGTAGGAACCTTATTCTAGCTCGATATGCGCTAGCAATTAGAA----- 2452
Qy 608 AsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThr 627
Db 2453 ---AAAGCAGCGGAAATATCGTGTAAAGACTCAGATATGCTACTGATGCA----- 2497
Qy 628 LeuProAsnIleSerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsn 647
Db 2498 -----GATATTGTATTGCATGTAACGATGCTCAGATTCAGATGCCAAAAACAATGAAC 2551
Qy 648 AsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGln 667
Db 2552 CCAGGT-----GAGGATCTGCATCTTAAACTTTTAAAGTTGCGAGAT 2593
Qy 668 PheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAla 687
Db 2594 GCTATCACAACTAAATTTAGCAACAGATAGTTCGCTAGCAATTGAAACAATAATTTAGGT 2653
Qy 688 AspValSerAsnSer-----IleLeuIleIleAspIysIleGluPheIlePro 703
Db 2654 GAAGACCTCAATTAATCAACATTTATCTGTATAGTTTACGTTGACCAATTCGAAATTCATCCCA 2713
Qy 704 IleThrSerSerMethIleGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLys 723
Db 2714 GTAGATGAGACATAT-----GAAGCGGACACAGATTTAGACGACGAGAAAGCA 2764
Qy 724 IleAsnThrPhePheThrAsn 730
Db 2765 GTGAATGCCTTGTTCAGGAAT 2785

RESULT 11
US-11-108-389-17
; Sequence 17, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 4874
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Genomic DNA 1218-1
; US-11-108-389-17

```


Alignment Scores:			
Pred. No.:	3.51e-61	Length:	4874
Score:	675.00	Matches:	216
Percent Similarity:	44.46%	Conservative:	125
Best Local Similarity:	28.16%	Mismatches:	300
Query Match:	17.45%	Indels:	126
DB:	7	Gaps:	32
US-10-783-417-2 (1-735) x US-11-108-389-17 (1-4874)			
QY	5	AsnAspAsnAsnGluTyrGluIleLeuAspSerHisThrSerProTyrPheProAsnArg	24
DB	740	AAATATCAAAATGAATATGAATATATAGATCGACACTTCTT-----ACTTCT	787
QY	25	AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn	43
DB	788	GTATCCAAATGATCTAACAGATACCCCTTTTCGGATGAGCCACAAATCGCTACAAAT	847
QY	44	ThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn	62
DB	848	ATGGATTATAAGATTATTTAAAAATGTCTCGGGAAATCTAGTGAATACCTCGTGTCA	907
QY	63	PheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIleValSer	82
DB	908	CCTGAAGTACTGTTACGGCAAGAT-----GCAGTAAGCGCCGAATGATATAGTA	961
QY	83	GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIle	102
DB	962	GGTAATAATTACTATCAGGTTTAGGG-----GTCCCAATTTGTT	997
QY	103	GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGlyGluGln	122
DB	998	GGCCGATAGTAGTCTTTATATCTCAACTATTGTATATCTGTGCGCTTCAGGGGAA	1054
QY	123	AspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu	142
DB	1055	---AAGAGTCATGGGAAATTTTATGGAACAGTAGAGAACTCATTTATCAAAAAATA	1111
QY	143	ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln	162
DB	1112	GCAGAATATGCAAGGAATAAAGCGCTTTCGGAATTAGAAGATTAGGTAATAATTACCA	1171
QY	163	SerTyrAsnThrAlaLeuAspAspTyrPArgLysLeuLysArgLeuGlnAlaProGlyLeu	182
DB	1172	TTATATCTAACTCGCTTGAAGATGCGGAAGAA-----1204	1204
QY	183	ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal	202
DB	1205	AAATCCAAATGGTCA-----AGAGCCCTTACGAGATGTGGAAATCGATTTGAATCCTG	1258
QY	203	HisAsnAspPheLeuArgGluIleProGlyPheGlnLeuGluTyrLysThrLeuLeu	222
DB	1259	GATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTACCATTC	1318
QY	223	LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlyAlaGlu	242
DB	1319	CTTACTGTATATGCAATGGCAGGCCAACCTTCATTACTGTATTAAAGGACGGGTCAATT	1378
QY	243	LeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr	262
DB	1379	TTTGGAGAAAGTGGGATGG-----TCACAACTACT	1411
QY	263	SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla	282
DB	1412	ATTAATAACTATTATGATCGTCAAAATGAAATCTTACTCGAGAAATTTTGATCACTGTGA	1471
QY	283	AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle	302
DB	1472	AAGTGTATGAAACTGTGTTAGCMAAATTAAGGCGCAGCGCTAAACCAATGGTGTGAC	1531
QY	303	PheAsnAspTyrArgArgTyrMetThrIleThrValLeuLeuAspThrIleSerGlnPheSer	322


```
QY 648 AsnThrPheSerGlyThrAsnThrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheCln 667
Db 2552 CCAGGT-----GAGATCTGACATCTAAACTTTTAAAGTTGCGAGAT 2593
QY 668 PheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAla 687
Db 2594 GCTATCACAACTTAATTTAGCAACAGATAGTTTCGTAGCATGAAACATAATTTAGGT 2653
QY 688 AspValSerAsnSer-----IleLeuIleIleAspLysIleGluPheIlePro 703
Db 2654 GAAGACCTAATTCACATTCGTGTATAGTTTACGTTGACCGAATCGAATTCATCCCA 2713
QY 704 IleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLys 723
Db 2714 GTAGTCAGACATAT-----GAAGCGGAACAGATTTAGAACGACGCGAAGAAAGCA 2764
QY 724 IleAsnThrPhePheThrAsn 730
Db 2765 GTGAATGCCTTGTGTACGAAT 2785

RESULT 12
US-11-091-643-3
; Sequence 3, Application US/11091643
; Publication No. US20050246789A1
; GENERAL INFORMATION:
; APPLICANT: TANAKA, Masao
; APPLICANT: YOKOYAMA, Tomoko
; APPLICANT: AOYAGI, Moriochi
; APPLICANT: HASEGAWA, Makoto
; APPLICANT: EHARA, Gaku
; APPLICANT: KIMURA, Masaharu
; APPLICANT: NISHIHASHI, Hideji
; TITLE OF INVENTION: Polypeptide having larvae growth inhibiting or
; TITLE OF INVENTION: insecticidal effect on scarabaeidae insects and
; FILE REFERENCE: OP1335
; CURRENT APPLICATION NUMBER: US/11/091,643
; CURRENT FILING DATE: 2005-03-29
; PRIOR APPLICATION NUMBER: JP 2001-115754
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: JP 2001-203463
; PRIOR FILING DATE: 2001-07-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 4359
; TYPE: DNA
; ORGANISM: Bacillus popilliae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (282)..(4229)
US-11-091-643-3

Alignment Scores:
Pred. No.: 1,3e-60 Length: 4359
Score: 669.00 Matches: 230
Percent Similarity: 43.47% Conservative: 106
Best Local Similarity: 29.75% Mismatches: 301
Query Match: 17.29% Indels: 136
DB: 9 Gaps: 34

US-10-783-417-2 (1-735) x US-11-091-643-3 (1-4359)

QY 1 MetAsnGln-----AsnAsnAspAsn-----AsnGluTyr 10
Db 336 ATCAATCAATATCATACCAAAACGATACAAAAGTTTACAAACCAAGTGGAAATGAAATG 395
QY 11 GluIleIle-----AspSerHisThrSerProTyrPheProAsnArgAsnSerAsnAsp 28
Db 396 CAATCATTCACCTTCAGTACGCTTTACTTTACAGTCCCAT----- 440
QY 29 SerArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsnThrAsnTyrLysGlu 48
```

```
Db 441 ---AAGTATCCGTATGTACCGATCCCAATGTCTACGAGCGGTGGAAAGTTATATAAAT 497
QY 49 TrpLeuAsnMetCysGlnGlyAsnThrGlnTyrGlyAspAsnPheGluThrPheAlaSer 68
Db 498 TGGTTGGATATGTGTACAGGACA-----GGCGACACACAGAAAGCCCGAAACTGTCT 548
QY 69 AlaAspThrIleAlaIaValSerAlaGlyThrIleValSerGlyThrLeuLeuAlaGly 88
Db 549 GCTATTTCAAAGGTGCTTTTCTGCTGCAATTAATAAGC----- 590
QY 89 IleGlyGlyLeuThrSerIleSerGlyProIleGlyIle-----IleGly 103
Db 591 -----ACGGGCTTCTGGCTTACTAGGTTCCTGTTGCCATCAAAATCGGG 638
QY 104 AlaIleIleSerPheGlyThrLeuLeuThrValPheThrProAlaGlyGluGlnAsp 123
Db 639 GCATTT-----TATACCTTCTTATTAATACCTTATGGCCTGCA-----AGC 680
QY 124 LysThrValThrThrGlnPheIleIleMetGlyGluIlePheValAspThrProLeuThr 143
Db 681 AATACTCAATGGGAGCAGTTTATAGCACATGTGGAAGAACTCATAAATGCAAACTAACA 740
QY 144 GluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGlnSer 163
Db 741 GATCATGTAGAATAATTCGGCACTTACAAAATTAATGTTTACGTAATAACATAGAGATA 800
QY 164 TyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaProGlyLeuPro 183
Db 801 TATAACGAAGCTTTAATAGTTTGGAAACAA-----GATCCT 836
QY 184 ProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHis 203
Db 837 AACAAATAGCAAACTAAAGAT-----CATGTAGCAAGTAAATTCGTAGTCTAAAT 887
QY 204 AsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeuLeu 223
Db 888 AGCCAATTCGAAGAATATATATTCACAAATTTAAGAAGAAAGTTTGAAGTTCAAATTA 947
QY 224 ProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeu 243
Db 948 ACTATATATGCACAACTCTGCAATCTTCATCTATTTATTAAGAGATTCCTCTTTGAT 1007
QY 244 AlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSer 263
Db 1008 GGTGCATCTTGGGATTT-----GCTCAAGCTACTATT 1040
QY 264 AspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAlaAsn 283
Db 1041 GACAAATAATTACATTCGCCAAATAAGGAAACCCGACAGATATGCAAAATCATTTGACCA 1100
QY 284 ThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysThrSerIlePhe 303
Db 1101 TGGTATCAGACGGGTTTACAAAGATTGCAAGGCACTCTGTCAGCATGTGGCTCTCTTAT 1160
QY 304 AsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeu 323
Db 1161 CATAGATTAGACAGAGAATGACACTAACAGTATTTGGATATTTGGCGCATTTATTTCAAAT 1220
QY 324 TyrAspIleLysArgTyrArgAspSerIleGlyIleGluLysGlyIleLysAsn 343
Db 1221 TATGATGCCGTAGTTAC-----CCACTGGAGGTAAAGGGA----- 1256
QY 344 GluLeuThrArgGluIleTyrThrThrGluIleAsn-----PheAspArg 358
Db 1257 GAGTCTACGAGAGAAATTTATACGATCCCGATCCCGATCTAACTGAGATATGATCGA 1316
QY 359 LeuProGlnLeuArgValGlnProAsnLeuAla---ThrMetGluTyrAsnLeuThrArg 377
Db 1317 GCACCATCATTCGAGAAATAGAAAATCTAGTAATTAGGGCACCAGAACTGTTACTTGG 1376
QY 378 AlaSerPheLysLeuPheSerPheGluGlnPheIlePheThrGluAsnThrAsn 397
Db 378 AlaSerPheLysLeuPheSerPheGluGlnPheIlePheThrGluAsnThrAsn 397
```

```
Db 1377 ATATCCGGTCAATTAACTAGTATATACAGGTAGATTGTACGGCTATACCTGGTAATAACGAT 1436
Qy 398 Phe-----GlyAsnArgLeuValGlyIleSerAsnArg-----AspAlaPro 411
Db 1437 TATTGGGCACACACAGGCTAGATTTCTTTGAAACCAATGTTATCGTTTGGGGTCTT 1496
Qy 412 ThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySerProThrThr 431
Db 1497 ACCTATGGATCGACGATTATATAAGT-----CGTACAGATTTCTATTTCCCATG 1544
Qy 432 LysThrIleArgProPheGluSerTyrLysValSerIleValThrAspArgGlnSerPro 451
Db 1545 AATTCATTT-----GATGTTATTCACACTACTAGTACTGTGGCTCTGCTTGG 1595
Qy 452 ProValSerProIleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGly 471
Db 1596 CCAACTGGC-----GGTTTGTGTTGGAGTCGCTTCGGCTAGATTTTTCG 1643
Qy 472 SerSerAsnAsnThrLeuLysTyrSerAlaGlyGlySerLeuSerAsn-----TyrGln 489
Db 1644 AAAAGTCCTAGCACC-----GGTTTATTAGTGAGCGGGTGTATCAG 1685
Qy 490 AsnThrThrPhePheGlnPheProArgLysLysAspCysAsnLeuValIleAsp---Pro 508
Db 1686 AATCCAGTATATTTT-----TCGAGTTCACCTTTAACTTTAACTTACCT 1730
Qy 509 GlyCys-----SerProAsnPheAsnAsnTyrSerHisIleLeuSerHisPheSer 525
Db 1731 GAGTAGACCAACATAGCCGACTGTCGCGACTATAGTATTAACATTCGTGTATCACA 1790
Qy 526 LeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeu 545
Db 1791 GCATTTTCGA-----ACTGGATTGAATGGAACCTGTT-----CCGGTITTT 1829
Qy 546 GlyTrpThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMet 565
Db 1830 GGACGGTATTCTGCACACTGTAGTCGACATAGAAATTCGAGCAGACAAATAACGCAA 1889
Qy 566 IleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGluGlyProGly 585
Db 1890 ATCCCGCGTGTAAAGCAAACTCCCTCGAC---AATGTCCAGTAGTTAGAGGACTGGA 1946
Qy 586 HisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIleThrCysGlu 605
Db 1947 TTTACAGGAGGACGCTGTTGAAGACAGTATCTTAGTGTGTTTTCCTCACTATCACT 2006
Qy 606 ThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThr-----621
Db 2007 TCATCGAGAGCGGGCCAAATCTTACCGCATCCGCTCGCTTATGCTGCTGAGTAGATTTA 2066
Qy 622 ---AsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGly 640
Db 2067 ATTATGAGTATATTTCTAATGACCTTTTATTTCCAAAGAAATAGCTTTACCAATCA 2126
Qy 641 IleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyr 660
Db 2127 ATGCCACCA-----CTGACCGAACT-----GTACCTACGAGCTTTTAATTT 2171
Qy 661 GlyAspPheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIle 680
Db 2172 GCAGATTTTGGTGC-----ACTTTTACAACAGCTACTGCTAACAAGATAT 2219
Qy 681 ProPheIlePheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIleGlu 700
Db 2220 ACTTTTCAATTCAT-----ACGGGTGGAGCAGCTATAATGACAGAAATGAA 2267
Qy 701 PheIleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIle 720
Db 2268 TTGTTCCAAATTGAGGGTAGTTTGTTCGAGTAGCAACCAACACACAGCTAGAAAAGCA 2327
Qy 721 GlnThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db 2328 AGGAAAGCGGTGAACCATTTGTTTACAGATGGATCGAAA 2366
```

RESULT 13

```
US-11-058-727-73
; Sequence '73, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnall
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins with Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 2025
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (mutated)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2025)
US-11-058-727-73
```

Alignment Scores:

```
Pred. No.: 6,45e-60 Length: 2025
Score: 657.50 Matches: 211
Percent Similarity: 44.67% Conservative: 120
Best Local Similarity: 28.48% Mismatches: 293
Query Match: 16.99% Indels: 117
DB: 31 Gaps: 31
```

US-10-783-417-2 (1-735) x US-11-058-727-73 (1-2025)

```
Qy 5 AsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyrPheProAsnArg 24
Db 10 AATAATCAAAATCAATATATGAATATATAGATCGCACCTTCT-----ACTTCT 57
Qy 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
Db 58 GTATCCAATGATTCTTAACAGATACCTTTTTCGGAATGAGCCACAAATGCGCTACAAAT 117
Qy 44 ThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
Db 118 ATGATTATTAAGATTATTTTAAATAATGCTCGGGAAATGCTAGTAATACCTCGTTCA 177
Qy 63 PheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSer 82
Db 178 CCTGAAGTACTTGTAGCGGACAAGAT-----GCAGTAAGCGCGCAATTGATATAGTA 231
Qy 83 GlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
Db 232 GGTAAATTAATCTACTCAGGTTTAGGG-----GTCCCATTTGTT 267
Qy 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGlyGluGln 122
Db 268 GGCGCCGATAGTAGTCTTTTATCTCAACTTATTGATATTCTGTGGCCTTCAGGGAA--- 324
Qy 123 AspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu 142
```

Db 325 ---AAGAGTCAATGGGAAATTTTATGGACAAGTAGAAGAACTCATTAAATCAAAAAATA 381
Qy 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
Db 382 GCAGAAATATGCAAGAAATAAAGCCCTTTCCGAATTTAGAAAGATTAGTAAATAATTACCA 441
Qy 163 SerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
Db 442 TTATATCTAATCGCTGCTGAAGATGGGAAGAA---AATCCATTCGAAGTCGAGTGT 498
Qy 183 ProProSerSerAlaLeuGlnGlnAlaLeuThrLeuLysIleArgPheGluAsnVal 202
Db 499 CGAAGTCGAGTGCCTTACGAGAT-----GTGCAAAATCGATTTGAAATCCTG 546
Qy 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
Db 547 GATAGTTTATTTAGCCAAATATATCCATCTTTTAGAGTGACAAAATTTTGAAGTACCAATC 606
Qy 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGlyAlaGlu 242
Db 607 CTYACTGTATATGCATGGCAGCCACCTTCATTTACTGTATTAAAGGACGCGTCAATT 666
Qy 243 LeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
Db 667 TTTGGAGAAGATGGGATGG-----TCAACAACACTACT 699
Qy 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
Db 700 ATTAATAACTATTATGATCGTCAAAATGAACCTTACTGCAGAAATATCTGTACACTGTGTA 759
Qy 283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTyrSerIle 302
Db 760 AAGTGGTATGAACCTGTTTAGCAAAATTTAAAGGACGACGCGCTAAACAATGGGTGGAC 819
Qy 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db 820 TATTAACCAATTCCTAGAGAAATGACACTGGCGGTTTTAGATGTGTGTCAATTATCCCA 879
Qy 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyIleGluValLysGlyIleLys 342
Db 880 AATTATGACACACGACGCTAC-----CCAATGGAAACGAAAGCA----- 918
Qy 343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362
Db 919 ---CAACTAACAGGGAAGTATATACA-----GATCCACTGGCGGCGGTA 960
Qy 363 ArgValGln-----ProAsnLeuAlaAlaThrMetGluTyr 373
Db 961 AACGTGCTTCAATGGTTCTCGTATGACAAAGCACCTTCTTCGGAGTGATAGAATCA 1020
Qy 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
Db 1021 TCCGTTATTCGACCACCC---CATGTATTGATTATATAACGGGACTCAGAGTGATACA 1077
Qy 394 GluAsnThrAsnPheGlyAsn---ArgLeuVal-----GlyIleSer 406
Db 1078 CAATCAAGAAGCACTTCTTCGCTCGCTATATAAGACATTGGGCTGGTGCATCAATAAGC 1137
Qy 407 AsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThr 426
Db 1138 TACCATCGTGTAGTAGGGGTAGTAAT-----CTTCAACAAATGTATGGAACCTAATCA 1191
Qy 427 GlySerProThrThrLysThrIleArgProPheGluSerTyrLysValSerIleValThr 446
Db 1192 AATCTACAGCAGCTAGTACCTTT---GATTTACGAATATGATATTATACAGACTCTA 1248
Qy 447 AspArgGlnSerProProValSerProIleGlnPro-----HisPheIle 461
Db 1249 TCAAGAGTGCAGTACTCTTGATATTGTTTACCTGTTATACGTATATATTTTGG 1308
Qy 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
Db 1309 ATGCCAGAAGTCGAGTGTTCATCGTTAAACCAATTTGAATATACCAAGAACGTTAAAG 1368

Qy 479 TyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArg 498
Db 1369 TAT-----AATCCAGTTTCCAAAGATATTATTATACCGAGTACAGA 1407
Qy 499 LysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSer 518
Db 1408 GATTCGGAAATTAGATTACCTCCAGAACTTCAGATCAACCAANTATCAGTCATATAGC 1467
Qy 519 HisIleLeuSerHisPhe---SerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeu 537
Db 1468 CATAGATTATGTCATATACAAAGTATTCCCGCGACGGGTAAACACTACCGGATTA----- 1521
Qy 538 GlnIleLeuAspThrGlyValLeuGlyTyrPheHisSerSerValAspArgTyrAsnAla 557
Db 1522 -----GTACCTGTATTTCTTGGACACATCGAAGTGCAGATTAAACAATACATA 1569
Qy 558 IleSerAspLysIleIleThrMetIleProAlaIleLys---GlyAsnAsnLeuAspThr 576
Db 1570 ATATATTCAGATAAATCACTCAAAATTCGGCCGTTAAATGTGGGATTAATTTACCGTTT 1629
Qy 577 AsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuVal---TyrLeuGln 595
Db 1630 ---GTTCCAGTGTAAAGGACGACACATACAGAGGGGATTATTACAGTATAATAGA 1686
Qy 596 SerGlnGly-----ArgLeuGluIleThrCysGluThrPro 607
Db 1687 AGTACTGGTCTGTAGGAACCTTATTCTAGCTCGATATAGGCTAGCATTAGAA----- 1740
Qy 608 AsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThr 627
Db 1741 ---AAAGCAGGAAATATCGTGTAAAGACTGAGATATGCTACTGATGCA----- 1785
Qy 628 LeuProAsnIleSerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsn 647
Db 1786 -----GATATGTATTGCTATGATAACGATGCTCAGATTCAGATGCCAAAAACAATGAAC 1839
Qy 648 AsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGln 667
Db 1840 CCAGGT-----GAGGATCTGACATCTAAAACTTTTAAAGTTGCAGAT 1881
Qy 668 PheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAla 687
Db 1882 GCTATCAACAGGTAAATTTAGCAACAGATAGTTCGGTTGCGAGTTAAACATATTTAGGT 1941
Qy 688 AspValSerAsnSer-----IleLeuIleIleAspLysIleGluPheIlePro 703
Db 1942 GAAGACCTTAATTCACATTATCTGTTATAGTTTACGTTGACCGAATCAATTCATCCCA 2001
Qy 704 Ile 704
Db 2002 GTA 2004

RESULT 14

US-11-058-727-79

; Sequence 79, Application US/11058727

; Publication No. US20050261483A1

; GENERAL INFORMATION:

; APPLICANT: Andre R. Abad

; APPLICANT: Ronald D. Flannagan

; APPLICANT: Rafael Herzmann

; APPLICANT: Theodore W. Kahn

; APPLICANT: Albert L. Lu

; APPLICANT: Billy Fred McCutchen

; APPLICANT: James K. Presnall

; APPLICANT: James F.H. Wong

; APPLICANT: Cao-Guo Yu

; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal

; TITLE OF INVENTION: Activity

; FILE REFERENCE: 35718/287809

; CURRENT APPLICATION NUMBER: US/11/058,727

; CURRENT FILING DATE: 2005-02-15

; PRIOR APPLICATION NUMBER: 60/391,786

```

; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 79
; LENGTH: 2025
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (mutated)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2025)
; US-11-058-727-79

Alignment Scores:
Pred. No.: 6,45e-60 Length: 2025
Score: 657.50 Matches: 212
Percent Similarity: 44.80% Conservative: 120
Best Local Similarity: 28.61% Mismatches: 292
Query Match: 16.99% Indels: 117
DB: 31 Gaps: 31

US-10-783-417-2 (1-735) x US-11-058-727-79 (1-2025)

Qy 5 AsnAspAsnGluThrGluLeuIleAspSerHisThrSerProTyrPheProAsnArg 24
Db 10 AATAATCAAAATGAATATGAATATATAGATCGCACACCTTCT-----ACTTCT 57
Qy 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
Db 58 GTATCCAAATGATCTTAACAGATACCCCTTTTGGCAATGAGCCAAACAAATGGCTTACAAAAT 117
Qy 44 ThrAsnTyrLysGluTrpLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
Db 118 ATGGATTATAAAGATATTATTAATAAATGTCTCGGGAATAGCTAGTGAATACCTCGTTCA 177
Qy 63 PheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSer 82
Db 178 CTGAAGTACTTCTTAGCGGACCAAGAT-----GCAGCTAAGCGCGCAATTGATAGTA 231
Qy 83 GlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
Db 232 GGTAAATTAATCTACAGGTTTAGG-----GTCCCAATTTGTT 267
Qy 103 GlyAlaIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGln 122
Db 268 GGGCGATAGTAGTCTTTTACTCAACTATTGATATTCTGTGGCTTCAGGGAA--- 324
Qy 123 AspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu 142
Db 325 ---AAGAGTCATGGGAATTTTATGGAAACAAGTAGAAGAACTCATTAATCAAAAATA 381
Qy 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
Db 382 GCAGATATCAAGGAATAAAGCGCTTCGGAATTAGAAGATTAGTAATAATTACCAA 441
Qy 163 SerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
Db 442 TTATATCTAACTCGCTTGAAGAATGGGAAGAA---AATCCATTTCCGAATCGAGGTCCA 498
Qy 183 ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal 202
Db 499 ATGGTTCCTCGGCGCTTACAGAT-----GTGCGAATCGATTGAAATCCCTG 546
Qy 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
Db 547 GATAGTTATTTACGCAATATATGCCATCTTTAGAGTAGCAAAATTTGAAGTACCATTC 606
Qy 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlyAlaGlu 242
Db 607 CTTACTGTATGCAATGGCAGCAACCTTTCAATTACTGTATTATTAAGGAGCGGCTCAATT 666

```

```

Qy 243 LeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
Db 667 TTTGGAGAGAAATGGGGATGG-----TCAACAACACTACT 699
Qy 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
Db 700 ATTAAATCAATATTATGATCGTCAATGAACTTACTCGAGAAATATTTGATCAGTGTGTA 759
Qy 283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle 302
Db 760 AAGTGGTATGAAACTGTTTAGCAAAATTAAGAGCGACGCGCTAAACAATGGTTCAC 819
Qy 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db 820 TATAACCAATTCGCTAGAGAAATGACACTGGCGGTTTGTAGATGTTGTTCATTATCCCA 879
Qy 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIleLys 342
Db 880 AATTATGACACACGCGACGTAC-----CCAATGGAAACGAAAGCA----- 918
Qy 343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362
Db 919 ---CAACTAACAGGGAAGTATATACA-----GATCCACTGGGCGCGTA 960
Qy 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
Db 961 AACGTGCTTCAATTGGTTCCTGCTATGACAAAGCACCTTCTTCGGAGTATAGATACTA 1020
Qy 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
Db 1021 TCGGTTATTTCGACACCC-----CATGTATTGATATATATACGGGACTCACAGTGTATACA 1077
Qy 394 GluAsnThrAsnPheGlyAsn---ArgLeuVal-----GlyIleSer 406
Db 1078 CAATCAAGAGCATTTCTTCGCTCGTATATATAGACATTTGGGCTGTCATCAATAAGC 1137
Qy 407 AsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThr 426
Db 1138 TACCATCGTGTAGTAGGGTAGTAAT-----CTTCAACAAATGATGGAACATAATCAA 1191
Qy 427 GlySerProThrThrLysThrIleArgProPheGluSerTyrLysValSerIleValThr 446
Db 1192 AATCTACACAGCACTAGTACCTTT---GATTTTACGAATTTATGATATTTTACAAGACTCTA 1248
Qy 447 AspArgGlnSerProProValSerProIleGlnPro-----HisPheIle 461
Db 1249 TCAAGATCGAGTACTCTCTTGATATTGTTTACCCTGTTATACGTATATATTTTGGGA 1308
Qy 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
Db 1309 ATCCAGAACTCGAGTTTTCATGGTAAACCAATTTGAATAATACCAGAAAGACGTAAAG 1368
Qy 479 TyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArg 498
Db 1369 TAT-----AATCCAGTTTCCAAAGATATATTAGCGAGTACAAGA 1407
Qy 499 LysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSer 518
Db 1408 GATTTCGAATTAGAATTACTCCAGAAATTCAGATCAACCAAAATTTATGAGTCATATAGC 1467
Qy 519 HisIleLeuSerHisPhe---SerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeu 537
Db 1468 CATAGATTATGTCATATACAGATTTTCCCGCAGCGGTAACACTACCGGATTA----- 1521
Qy 538 GlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAla 557
Db 1522 -----GTACCTGTATTTCTTGACACATCGAAGTGCAGATTTTAAACAATACA 1569
Qy 558 IleSerAspLysIleIleThrMetIleProAlaIleLys---GlyAsnAsnLeuAspThr 576
Db 1570 ATATATTAGATAAAATCAATCAAAATCCGCGCGTTAAATGTTGGGATAAATTTACCGTTT 1629

```

```
QY 577 AsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuVal---TyrLeuGln 595
Db 1630 ---GTTCCAGTGGTAAAGGACCGAGCATACAGAGGGGATTTATTACAGTATAATAGA 1686
QY 596 SerGlnGly-----ArgLeuGluIleThrCysGluThrPro 607
Db 1687 AGTACTGGTCTGTAGGAACCTTATTTCTAGCTCGATATGCCCTAGCATTAGAA----- 1740
QY 608 AsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThr 627
Db 1741 ---AAAGCAGGGAATATCGTGAAGACTGAGATATGCTACTGATGCA----- 1785
QY 628 LeuProAsnIleSerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsn 647
Db 1786 -----GATATTGTAATGCGATGAACGATGCTCAGATTGAGTGCACAAAACATGAAC 1839
QY 648 AsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGln 667
Db 1840 CCAGGT-----GAGGATCTGACATCTAAACCTTTTAAAGTTGCAGAT 1881
QY 668 PheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAla 687
Db 1882 GCTATCACACAGTAAATTTAGCAACAGATAGTTTCGGTTCAGTTAAACATAATTTAGGT 1941
QY 688 AspValSerAsnSer-----IleLeuIleIleAspLysIleGluPheIlePro 703
Db 1942 GAAGACCTTAATCAACATATCTGTTATAGTTTACGTTGACCAATCGAATTCATCCCA 2001
QY 704 Ile 704
Db 2002 GTA 2004
```

RESULT 15

```
US-11-108-389-73
; Sequence 73, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins with Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 2025
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (mutated)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2025)
US-11-108-389-73
```

Alignment Scores:	
Pred. No.:	6.45e-60
Score:	657.50
Percent Similarity:	44.67%
Best Local Similarity:	28.48%
Length:	2025
Matches:	211
Conservative:	120
Mismatches:	293

```
Query Match: 16.99% Indels: 117
DB: 7 Gaps: 31
US-10-783-417-2 (1-735) x US-11-108-389-73 (1-2025)
QY 5 AsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyrPheProAsnArg 24
Db 10 AATTAATCAAAATGAATATGAATATATAGATCGGACACCTTCT-----ACTTCT 57
QY 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
Db 58 GTATCCAAATGATTTAAACAGATACCTTTTCGGAATGAGCAACAATCGCTACAAAT 117
QY 44 ThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
Db 118 ATGATTATTAAGATATTTAAATAATGTCTCGGGAATGTAGTGAATACCTCGTTCA 177
QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIleValSer 82
Db 178 CCGAGTACTTGTTCGCGCAAGAT-----GCAGCTAAGCGCGCAATTCATATAGTA 231
QY 83 GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
Db 232 GGTAAATTAATCTATCAGGTTTAGGG-----GTCCCATTTGTT 267
QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGlyGln 122
Db 268 GGGCGGATAGTAGCTCTTTATCTCACTTATCTGTCGCGCTTCAGGGGAA--- 324
QY 123 AspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu 142
Db 325 ---AAGAGTCAATCGGAAATTTTATGGAAACAAGTAGAAGAACTCATTAATCAAAAAA 381
QY 143 ThrGluSerIleLysGlnLeuLysGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
Db 382 GCAGAAATGCAAGGAATAAAGCGCTTTCGGAATTAGAGGATAGGTAGTAATTAATACCA 441
QY 163 SerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
Db 442 TTATATCTAATCGGCTTGAAGATGGAAGAA---AATCCATTTTGAAGTCGAGGTTT 498
QY 183 ProProSerAlaLeuGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal 202
Db 499 CGAAGTCGAGTGCTTACGAGAT-----GTGCAAAATCGCAATTTGAATCCTG 546
QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
Db 547 GATAGTTTATTTAGCAATATATGCCATCTTTAGAGTGACAAATTTTGAAGTACCATTC 606
QY 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGlu 242
Db 607 CTTACTGTATATGCAATGGCAGCAACCTTCATTCTGTTATTAAAGAGCGCGTCAATT 666
QY 243 LeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
Db 667 TTTGGAGAAGAAATGGGGATGG-----TCAACAACACTACT 699
QY 263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
Db 700 ATTAATAACTATTATATGATCGTCAAAATGAACCTTACTGCAGAAATATCTGCATCCTGTGTA 759
QY 283 AsnThrTyrArgThrGlyLeuLysAsnLeuAspGluProAsnMetLysTrpSerIle 302
Db 760 AAGTGGTATGAACCTGGTTTAGCAAAATTAAGAGCGCAGCGCTAAACAATGGGTGAC 819
QY 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db 820 TATAACCAATTCCTAGAGAAATGACACTGGCGGTTTGTAGATGTGTGCTATTATCCCA 879
QY 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIleLys 342
Db 880 AATTATGACACACGACGCTAC-----CCAATGGAAACGAAAGCA----- 918
```

```
QY 343 AsnGluLeuThrArgGluIleTyrThrGluIleAsnPheAspArgLeuProGlnLeu 362
Db 919 ---CAACTAAAGGAAGTATATACA-----GATCCACTGGGCGCGGTA 960
QY 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
Db 961 AACGTGCTCAATTGGTTCCTCGTATGACAAAGCACCTTCTTCGGAGTGATAGATCA 1020
QY 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393
Db 1021 TCGGTTATTGACCAACCC---CATGTATTGATTATATAACGGGACTCACAGTGTATACA 1077
QY 394 GluAsnThrAsnPheGlyAsn---ArgLeuVal-----GlyIleSer 406
Db 1078 CAATCAAGAAGCAATTTCTCCGCTCATATATAAGACATTCGGCTGTCATCAAAATAAGC 1137
QY 407 AsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThr 426
Db 1138 TACCATCGTGTAGTAGGGTAGTAAAT-----CTTCAACAATGATGGAACTAATCAA 1191
QY 427 GlySerProThrThrLysThrIleArgProPheGluSerTyrLysValSerIleValThr 446
Db 1192 AATCTACACAGCAGCTAGTACTTT---GATTTTACGAATTATGATATTTACAAGACTCTA 1248
QY 447 AspArgGlnSerProProValSerProIleGlnPro-----HisPheIle 461
Db 1249 TCAAGAGGTGCAGTACTCTGTGATATTGTTTACCTCGTTATACGTATATATTTTGGGA 1308
QY 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
Db 1309 ATGCCAGAGTCGAGTTTTCATGGTGAACCAATTGAAATTAATCCGAAAGAGCGTTAAAG 1368
QY 479 TyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPheGlnPheProArg 498
Db 1369 TAT-----AATCCAGTTTCCAAAGATATTATAGCGAGTACAAGA 1407
QY 499 LysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSer 518
Db 1408 GATTCGGAATTAGAAATACCTCCGAAACTTCAGATCAACAAATTTAGTGCATATAGC 1467
QY 519 HisIleLeuSerHisPhe---SerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeu 537
Db 1468 CATAGATTATGTCATATAT---AAGTATTTCCCGCGCGGTAACTACCGGATTA----- 1521
QY 538 GlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAla 557
Db 1522 -----GTACTGTATTTCTTGACACATCGAAGTGCAGATTTAAACAATACA 1569
QY 558 IleSerAspLysIleThrMetIleProAlaIleLys---GlyAsnAsnLeuAspThr 576
Db 1570 ATATATTCAGATAAAATCACTCAAAATCCGGCGGTTAAATGTTGGGATATTTACCGTTT 1629
QY 577 AsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuVal---TyrLeuGln 595
Db 1630 ---GTTCCAGTGGTAAAGGACGAGCACATACAGGAGGGGATTTATTACAGTATATAGA 1686
QY 596 SerGlnGly-----ArgLeuGluIleThrCysGluThrPro 607
Db 1687 AGTACTGGTCTGTAGAACCTTATTCTTAGCTCGATATGGCTAGCAATTAGAA----- 1740
QY 608 AsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThr 627
Db 1741 ---AAAGCAGGGAATATCGTGTAGACTGAGATATGCTACTGTATGCA----- 1785
QY 628 LeuProAsnIleSerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsn 647
Db 1786 -----GATATTGTATTGCTAATGATAACGATGCTCAGATTCAGATTCGCAAAACAATGAAC 1839
QY 648 AsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGln 667
Db 1840 CCAAGGT-----GAGGATCTGACATCTAAACCTTTTAAAGTTGCAGAT 1881
QY 668 PheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAla 687
```

```
Db 1882 GCTATCAACACAGTTAATTTAGCAACAGATGTTTCGGTTGCAGTTAAACATAATTTAGGT 1941
QY 688 AspValSerAsnSer-----IleLeuIleIleAspLysIleGluPheIlePro 703
Db 1942 GAAGACCCCTAATTCACACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCA 2001
QY 704 Ile 704
Db 2002 GTA 2004
```

Search completed: December 11, 2005, 23:20:03
Job time : 252 secs